On a Controversy Regarding Primordial Gravitational Waves

Master thesis

Eirik Gjerløw
Institute for Theoretical Astrophysics
University of Oslo

June 1, 2010

Preface

The feature that perhaps unites most master theses is the disproportional amount of work put into it compared to its relevance for the rest of the scientific community. This text is no exception, but despite its irrelevance, pride and accomplishment are the feelings that dominates as the finishing touches are laid upon it.

As with most accomplishments, regardless of magnitude, the accomplishment of completing this text could not have happened were it not for a number of valuable and invaluable helpers:

Øystein Elgarøy: My thesis supervisor. He came up with the idea for this thesis, and has patiently answered all varieties of questions that have surfaced. I wish to thank him for investing time and confidence in me, and for always being supportive and helpful throughout the time it has taken to complete this thesis.

Hans Kristian Eriksen: Faithful mail-answerer and the oracle of all things statistical. He suggested many of the improvements to the likelihood analysis in chapter 9, and has also displayed an impressive amount of patience faced with my numerous questions. I wish to thank him for repeatedly helping me out, despite no legal obligation to do so.

Thomas Golding: Co-student and partner in procrastination. He has made the experience of writing a master thesis much more enjoyable than it would otherwise have been. I wish to thank him especially for proof-reading my thesis, which must have been pretty boring for a solar physicist as himself.

Dag Sverre Seljebotn: The local oracle on programming and statistics. I wish to thank him for being patient and always willing to answer questions.

Maren Grindstad: We have shared the yoke of Øystein's supervisorship together. I wish to thank her for many important discussions on the field of cosmology and the frustrations that are associated with working in it.

Staff and students at ITA: I wish to thank these for the excellent working environment they help create.

My teachers: This thesis is in a sense the end of a long path, and I would not have started walking it, nor continued walking, were it not for the teachers who have been there along the way. I wish to thank these for showing me that one can always know more about anything, and for cultivating my interest in the vast and fascinating subject that is physics.

Julie R. Gjerløw: My wife. I wish to thank her for always being supportive and kind, and for being patient when I have had to work late nights on this thesis.

I wish to thank my mother and father, the rest of my family, and all my friends.

Finally, I wish to thank my God, to whom everything is owed.

Contents

Chapter 1

Introduction

The field of cosmology in its current form can be said to trace its origins back to the theory of general relativity, proposed by Albert Einstein almost a hundred years ago. Since then, the various models of the Universe that have been proposed have either been strengthened or discarded on the basis of new observational data. There seems to consistently have been a surplus of models and demand of data, an instance of which was the Steady-State controversy, where two competing models of the Universe both seemed viable - until the blackbody spectrum of the Cosmic Microwave Background was observed. The last two decades have seen a radical increase in the amount of data, collected by such satellites as $COBE$ and $WMAP$, and by ground-based experiments. Nevertheless, many aspects of the current model of our Universe remain unclear - especially those that concern the early Universe.

In this text, one such aspect is investigated closely. The perturbed Einstein equations predict the existence of waves in the space-time fabric, called gravitational waves. The theory of inflation, or any theory relying on quantum physics to generate the initial conditions of the Universe, predicts that gravitational waves should be present in the early Universe, in addition to the density perturbations that eventually became all of the structure that we presently see. The theory of inflation predicts that if we quantify these two kinds of primordial perturbations - gravitational waves and density perturbations - the gravitational waves play a subdominant part.

There are, however, dissenting voices that do not agree with the predictions of inflationary theory, and one of these is that of Leonid P. Grishchuk. A pioneer in the field of gravitational waves, he started expressing his disagreement with the inflationary community in Grishchuk (1994), claiming that the inflationary predictions were both absurd and wrong. He still maintains this position, and coauthored Zhao et al. (2009a) and Zhao et al.

(2009b), in which analyses of the CMB are carried out, where the authors find stronger hints of gravitational waves than the standard $WMAP$ analysis does. It is claimed that this conflict stems from erroneous assumptions that the $WMAP$ team makes about the spectral index: A constant or simple running spectral index for all scales is wrong, and leads to wrong results.

The questions asked in this thesis are the following:

- Are Grishchuk's claims about the inflationary predictions (namely, that they are wrong) correct?
- Does the assumption of a constant spectral index lead to an erroneous negligence of signs of gravitational waves in the CMB?

In chapter 2, we start by reviewing some of the preliminary theoretical knowledge needed to understand the rest of the text. In Part I, we review some central parts of the Standard Model of cosmology: Chapter 3 reviews cosmological perturbation theory, chapter 4 reviews the theory of inflation, and chapters 5 and 6 review the theory behind the anisotropies in the CMB.

In Part II, we examine the various claims made by Grishchuk. Chapter 7 reviews the whole controversy that followed his 1994 paper, along with some of the pioneering work he has done in the field of gravitational waves. In chapter 8 we attempt to answer the first question above, by carefully analysing Grishchuk's theoretical derivations, while in chapter 9 we attempt to answer the second, by doing our own likelihood analyses of the $WMAP$ data.

Chapter 2

Preliminaries

2.1 The theory of General Relativity

Most of Part I review what has been dubbed the "Standard Model of cosmology", a highly successful (in terms of agreement with data) model of the Universe we inhabit. To be able to speak of any 'model' of cosmology, let alone anything we are permitted to call 'standard', we need a framework on which to build our models. For almost a hundred years, this framework has been the theory of General Relativity (henceforth abbreviated GR).

GR, in short, describes the interaction between spacetime and energy (which could be in the form of matter). It tells us that some matter distribution actually bends spacetime surrounding it, warping its properties. The mathematical framework of GR allows us to compute the effects of this bending and the trajectories of test objects in the vicinity of the distribution.

2.1.1 The metric and line element

Central to the theory of GR is the concept of a *metric*. The metric is described by a four-dimensional symmetric tensor of rank 2, called the *metric tensor*, conventionally labelled g. This quantity is arguably (and by 'arguably' we really mean: not arguable at all) the most central quantity in the theory of GR, and when solving any kind of dynamic problem in GR, one is certain to need to know the form of the metric tensor.

We can't just use whatever metric tensor that comes to mind. In particular, the metric tensor must be such that it satisfies what is called the *Einstein equations*.

The Einstein equations

The Einstein equations, in their simplest form, are:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{2.1}$$

Not so simple, we might say, since we know neither what $G_{\mu\nu}$ nor $T_{\mu\nu}$ are supposed to be - the scalar G in the equation above is Newton's gravitational constant.

The *tensor* G above is called the *Einstein tensor*, and it can be expressed as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R}, \tag{2.2}$$

which doesn't really clear things up - we now have the new quantities R and $\mathcal{R}$ instead. The tensor R is called the *Ricci tensor*, and the scalar $\mathcal{R}$ is the Ricci scalar, which is just the *contraction* of the Ricci tensor:

$$\mathcal{R} = R_{\mu\nu} g^{\mu\nu}. \tag{2.3}$$

This equation also defines what we mean by a contraction.

The Ricci tensor, in turn, can be expressed in different ways, but what will be useful here is to express it in terms of the *Christoffel symbol* Γ:

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\beta\alpha}\Gamma^{\beta}_{\mu\nu} - \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\mu\alpha}. \tag{2.4}$$

The Christoffel symbol is introduced in order to be able to define a *covariant derivative*: If, when differentiating a tensor, one only differentiates the tensor *components*, one misses the change of the tensor that arises because of the change of basis vectors that inevitably will occur because of a non-Cartesian space. Following such a differentiation procedure will render the new quantity that arises dependent on position - it becomes non-covariant. The Christoffel symbol incorporates the change of basis vectors between one point and another, and one way to define the Christoffel symbol is to define a covariant derivative (Grøn & Hervik, 2007)

$$A^{\mu}_{\ ;\nu} \equiv A^{\mu}_{\ ,\nu} + A^{\alpha}\Gamma^{\mu}_{\alpha\nu}, \tag{2.5}$$

where a comma denotes an ordinary derivative, and then demand that the quantity $A^{\mu}_{\ ;\nu}$ transform as a tensor. This defines the Christoffel symbol. Γ can be expressed in many ways, but limiting ourselves to a *coordinate basis*, where the basis vectors are taken to be derivatives: $\vec{e}_{\mu} = \frac{\partial}{\partial x^{\mu}}$, the one we will stick with is this:

$$\Gamma^{\mu}_{\alpha\beta} = \frac{g^{\mu\nu}}{2}\left(\frac{\partial g_{\alpha\nu}}{\partial x^{\beta}} + \frac{\partial g_{\beta\nu}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\nu}}\right). \tag{2.6}$$

Thus, the Christoffel symbol, and hence the Ricci tensor, Ricci scalar, and Einstein tensor, is completely determined by the metric tensor g.

The form of the *energy-momentum tensor* T is determined by the type of energy that fills the region of spacetime we are considering[1]. If that region is filled with a *perfect fluid*, that is, a fluid that has no viscosity or heat conduction (Grøn & Hervik, 2007), the energy-momentum tensor takes the simple form (Dodelson, 2003)

$$T^{\mu}_{\ \nu} = diag\{-\rho, p, p, p\}. \tag{2.7}$$

We will often in what follows assume the Universe to be filled with perfect fluids with various equations of state relating the energy density ρ to the isotropic pressure p through the simple law

$$p = w\rho, \tag{2.8}$$

where the parameter w determines the nature of the perfect fluid.

When the energy-momentum tensor is specified, we can make sure our metric tensor is a solution to the Einstein equations.

The line element

Knowing the actual form of the metric allows us to form a second very important quantity, called the *line element*, ds. Generally, when we know the components of the metric, the line element easily follows:

$$ds^2 = g_{\mu\nu} dx^{\mu} dx^{\nu}, \tag{2.9}$$

where dx^{μ} is an infitesimal interval between two spacetime events in the direction of the coordinate x^{μ}. Given two events, the line element is equal to *the time interval measured by an observer who is present at both events*: $ds^2 = -d\tau^2$, where τ is the *proper time* of the observer, proper time being defined by just the italicised sentence above.

The geodesic equation

The geodesic equation can be derived, for example, using the method of variations, and in symbols, it says

$$\frac{d^2 x^{\mu}}{d\lambda^2} = -\Gamma^{\mu}_{\alpha\beta} \frac{dx^{\alpha}}{d\lambda} \frac{dx^{\beta}}{d\lambda}, \tag{2.10}$$

[1] We will typically consider regions comparable to, or bigger than, the observable universe.

and describes the movement of some object[2] through spacetime. This equation is the GR equivalent of Newton's Second Law in the absence of external forces, with the left side being the acceleration of the object, and the right side involving the curvature of space, in some sense. The parameter λ must be such that it is monotonously increasing along the worldline of the object (for massive objects, $\lambda = \tau$ (the proper time of the object) is a practical choice). As we saw above, Γ contains information about the metric, and is therefore the link to GR.

2.2 Cosmology

The framework of GR presented in the previous section can be, and has been, applied to our Universe[3] as a whole in order to determine its evolution in spacetime. In this section, we will state some of the results of doing this. Our starting point will be to assume that our universe is homogeneous (it is the same no matter where we are) and isotropic (it looks the same in all directions).

2.2.1 The Friedmann-Robertson-Walker metric

The assumptions of homogeneity and isotropy alone lead (Grøn & Hervik, 2007) to a form of the metric known as the *Friedmann-Robertson-Walker (FRW)* metric. The line element for FRW universes is (Grøn & Hervik, 2007):

$$ds^2 = -dt^2 + a(t)^2\left(\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right). \tag{2.11}$$

Here, $a(t)$ is the *scale factor* which measures physical distances: The physical distance l from the origin to an object at radial coordinate r_{obj} is $l = a(t)\int_0^{r_{obj}} \frac{dr}{\sqrt{1-kr^2}}$. r, θ and ϕ, in turn, are the radial comoving[4] coordinates. The time parameter t is *cosmic time*, which is defined as the proper time of an observer who is at rest in the comoving coordinate system. The parameter k describes the geometry of the Universe: For $k = 1$, the spatial hypersurfaces of constant cosmic time has positive curvature - the Universe is closed. For $k = -1$, the hypersurfaces have negative curvature, leading

[2]This object could well be a light beam, or something else lacking mass.

[3]By 'our Universe' we will here, and subsequently, mean 'the observable universe'.

[4]Moving with the expansion of the universe, so that a stationary observer in these coordinates measures no net momentum density.

to an open Universe, while for $k = 0$, the hypersurfaces have Euclidean geometries, and the Universe is flat (Grøn & Hervik, 2007).

We will, for the rest of this text, work exclusively with a flat universe. Smoking hot observations (Komatsu et al., 2010) indicate that it is very close to flat indeed.

The FRW line element for flat universes follows directly from eq. (2.11) by setting $k = 0$:

$$ds^2 = -dt^2 + a^2(t)(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2). \tag{2.12}$$

2.2.2 Some useful quantities

We will here define, in rapid succession, some quantities that will be useful in the treatment that follows. We start with the definition of *conformal time* η:

$$\eta = \int_0^t \frac{dt'}{a(t')}. \tag{2.13}$$

This variable is a monotonically increasing parameter, which is convenient for us when dealing with general relativistic dynamics, as in eq. (2.10), where the parameter λ must have that trait. In other words, we can often use η as our monotonically increasing parameter in dynamic equations. From here on, overdots on a quantity will mean that the derivative with respect to conformal time is taken, while a prime means a derivative with respect to cosmic time t.

With our conventions (specifically, $c = 1$), η also becomes a distance. Indeed, it is a very significant distance: it is *the furthest any kind object could have travelled since the beginning of time*. Thus, no information may have travelled further than η.

Another useful quantity is the *Hubble parameter* $H = \frac{da/dt}{a}$. This parameter has a long and interesting history (Dodelson, 2003), but we shall mostly be concerned with its inverse $1/H$, which is called the *Hubble radius*, and which is often used to provide a rough estimate of the size of the observable Universe: during the course of one expansion time, that is, the time it takes for the scale factor to double, the Hubble radius is the distance over which particles can travel (Dodelson, 2003). So if two objects are separated by more than a Hubble radius, they cannot presently communicate. The comoving Hubble radius, $1/(Ha)$ is then a rough estimate of the comoving size of the Universe.

2.2.3 The Friedmann equations

We mentioned in section 2.1.1 that any metric has to satisfy the Einstein equations. However, we also stated that the assumptions of homogeneity and isotropy *alone* led to the FRW metric, eq. (2.11). This equation, though, does not completely specify the metric: We need to find the function $a(t)$, and in principle also the parameter k, for the metric to be completely determined.

Inserting the metric implied in eq. (2.11) into the Einstein equations, we eventually end up with the following *Friedmann equations*[5] (Grøn & Hervik, 2007):

$$(a')^2 + k = \frac{8\pi G}{3}\rho a^2, \tag{2.14}$$

$$a'' = -\frac{4\pi G}{3}(\rho + 3p)\, a. \tag{2.15}$$

Here, ρ is the total energy density, and p the total pressure of the Universe.

The Friedmann equations can also be written as

$$\Omega + \Omega_k = 1. \tag{2.16}$$

Here, $\Omega = \sum_i \Omega_i$, where $\Omega_i = \rho_i/\rho_c$ is the *density parameter* for fluid i. ρ_c is the *critical density*,

$$\rho_c = \frac{3H^2}{8\pi G}, \tag{2.17}$$

which is the total density the Universe must have to make it flat. The quantity Ω_k is a 'curvature density parameter', $\Omega_k = -k/(a^2H^2)$, which must be zero if the Universe is flat.

2.2.4 Solutions to the Friedmann equations

Analytical solutions to the Friedmann equations can only be found in certain special cases, and most easily if we assume that the Universe is filled with only one kind of fluid, so that we only have to deal with one equation of state. We will here simply list the solutions to the Friedmann equations for some well-known cases (all of them valid for flat universes only). We also show how each kind of energy density varies with a as this will help us understand the current model of the evolution of the Universe:

[5] It seems there are not complete unanimity on which, exactly, are the actual Friedmann equations; any manifestations of them can be reproduced with the equations presented here.

- Radiation-dominated universe ($w = 1/3$): $a \propto t^{1/2}$, $\rho \propto a^{-4}$. By radiation, we here mean all massless particles (such as photons and, to a good approximation, neutrinos) and particles whose energy is much higher than their rest energy.
- Dust (nonrelativistic matter)-dominated universe ($w = 0$): $a \propto t^{2/3}$, $\rho \propto a^{-3}$. 'Dust' is what we call all particles that are nonrelativistic (so that their energies are comparable to their rest energies), be they *baryons* (which, confusingly, in cosmology means both baryons and leptons) or so-called *cold dark matter* (matter that does not interact through the electromagnetic force).
- Cosmological constant-dominated universe ($w = -1$): $a \propto e^{Ht}$, $\rho = const.$ This solution is also called a *de Sitter* model.

In the next section, we shall see that each of these scenarios are valid in some part of the history of the Universe (at least according to the way that history is currently construed).

2.2.5 The Universe until recombination

The "Standard Model" (or, the ΛCDM model) of cosmology is the one currently dominant in the cosmological community. We need only mention the basic features of this model in order to make easier subsequent discussion, so this section will in no way contain a detailed exposition of the various stages of the evolution of the Universe. What follows is simply a crude list of these stages, and only until recombination, a concept explained below. All of the below, of course, is according to the standard ΛCDM model.

Inflation

A very short time after $t = 0$, the Universe, it is claimed, entered into a phase in which the energy density was dominated by scalar fields, giving it an effective equation of state $w = -1$. As we saw above, such an energy density leads to exponential expansion, and, as we shall see in chapter 4, this phase was the catalyst for creating the structure that we presently see in the Universe.

Radiation era ($t \approx 10^{-12}s$ - $t \approx 10^{4}yr$)

By some mechanism or other, inflation must have ended at some point. The Universe was then quite cold and dominated by nonrelativistic matter,

which after a while decayed into relativistic particles. This period is called *reheating*, and it ushered in the radiation era. The temperature during this era was so high that the photons and relativistic particles in the Universe constituted a single fluid. The concept of a universe expanding from an initial hot, radiation-dominated phase is called the "Hot Big Bang" (or just "Big Bang") model, and the ΛCDM model thus incorporates this concept, in addition to other concepts, like inflation.

Dust domination ($t \approx 10^4 yr$ - now)

We saw in the last section that the energy density of radiation has a higher rate of decrease as the scale factor increases than does the energy density of dust. As a result, there will be a time when the energy density of radiation is equal to the energy density of dust, and after this, the Universe is dust dominated. Its expansion is then slightly faster than during radiation domination. During this time, the radiation-matter fluid eventually cools down so much that electrons can combine with protons to form atoms - an event called *recombination*. After this, the photons are only coupled to the electrons, but as the electron fraction decreases rapidly, the photons decouple completely from matter and has since then (about 380,000 years after $t = 0$) streamed freely through the Universe, with the exception of the odd encounter with matter now and then. Later on, structure formed and became the Universe we know and love.

The above lists, very roughly, the various stages of the ΛCDM universe until recombination. It also seems (Becker et al., 2001; Fan et al., 2002) that the Universe has been reionised lately (an event called *reionisation*), so that electrons are again free to interact with the free photons. They won't form an equally tight plasma as before recombination, since the energy densities of both electrons and photons have decreased due to the expansion of the Universe, so that they interaction rate is much lower. This will nevertheless have an impact on the photons that pervade the Universe, as we shall see in chapter 6.

Denoting the contributions to the density parameter from dust, radiation and a cosmological constant by subscripts d, r, and Λ, respectively, we may state what cosmologists presently believe about the inventory of our Universe: We live in a flat ($\Omega_k = 0$) universe with $\Omega_r \approx 4 \times 10^{-5}$, $\Omega_d \approx 0.3$ and $\Omega_\Lambda \approx 0.7$. Moreover, they believe that most dust is in the form of cold dark matter, and only a small percentage of it ($\Omega_b = 0.04$) is in the form of baryons (Komatsu et al., 2010).

Part I

Standard Lore

Chapter 3

Not entirely smooth: The perturbed Universe

As mentioned in chapter 2, the FRW line element demands that the universe in question is both isotropic and homogeneous. Clearly, our universe is not entirely isotropic: The vista of my office desk is quite different from the view I get if I turn my head ninety degrees to face the window. Neither is it homogeneous: It *does* matter where in the Universe you happen to be situated (ask someone in the vicinity of a black hole, or inside a neutron star, whether it feels any different than at the surface of the earth).

However, my office and the scenery outside are *very* small in a cosmological context. In fact, even our Galaxy is small in that context, and it is only when we look at scales on the order of 300 million light years that space starts to look homogeneous. Luckily, the evolution of the Universe as a whole doesn't depend very much on what goes on at scales smaller than this, since gravity is a rather weak force.

Also, even though the Universe doesn't look exactly the same in all directions, it looks *roughly* the same - the amount of clumps and structure isn't noticeably larger in one part of the sky compared to the rest. In other words, the Universe is *statistically isotropic*.

All of the above means, in short, that the FRW line element is a good approximation to our universe. But it's not very exciting or informative to pretend that this approximation is the best we can do. For example, how 'clumpy' should the structure we observe be, if our theories are correct? Or, as pertains more to the current work: What about the small variations that we observe in the Cosmic Microwave Background - what are our predictions

for the amount and intensities of these? The FRW universe gives none. Therefore, we must try to make a better approximation to our Universe - and for that purpose, we use perturbation theory.

3.1 The basics of perturbation theory

When doing perturbation theory (at least in this context), we start out by assuming total homogeneity and isotropy in the Universe - i.e., we are assuming the situation to be exactly as described in section 2.2.1. We then add small perturbations to all relevant quantities; e.g. for the baryon density we let $\rho_b(t) \to \rho_b(t, \vec{x}) = \rho_b^{(0)}(t) + \delta\rho_b(t, \vec{x})$ where $\rho^{(0)}$ is the original (zero order) quantity (we will use this notation from now on), and $\delta\rho$ is the perturbation, which is allowed spatial as well as temporal variation. A 'small' perturbation is defined as "small enough that we can drop second-order terms and still maintain our preferred level of accuracy". This level of accuracy, a relatively subjective choice, then defines how large the perturbations can be before they become a liability to our calculations.

If we allow all perturbations to vary in space, we can then describe small deviations from isotropy and homogeneity. We have not, of course, said anything about how these perturbations arise in the first place - we will investigate that question in chapter 4.

We are, in the current work, mainly interested in the perturbations to the intensity and polarization distributions of the photons in the Universe. In order to find how these evolve, we need to use the Boltzmann equation - which is what we do in chapter 5.

Finally, in order to connect theory with observation, we must determine what effect the perturbations have on the things that we can observe. We will focus on one particular observable in chapter 6 - the CMB angular power spectrum.

Before we set off, we should make clear some things: As we mentioned above, we are mostly interested in the perturbations in the photon distribution. However, eq. (2.10), which holds for, say, a photon, shows that the way photons move depend on the shape of the metric at that space-time point. The metric, in turn, depends on matter (photons, baryons, dark matter) through eq. (2.1), so there is a complex interplay between photons, the metric, and the stuff that fills our Universe. Also, photons are not transparent to matter - they may collide with charged particles if these are present. In principle, then, in addition to photons, we need to find

equations governing both the metric perturbations and the perturbations in matter. However, in this text, we will limit ourselves to the treatment of photon perturbations, as we are mainly interested in showing *that* the perturbations set up during the early Universe influence the present photon distribution - not exactly *how* they do so (we *will* eventually compare the photon distributions that result from various cosmological models with each other, but that will be done with well-established software which takes into account all that is needed).

Metric perturbations hold a central position in much of this text, so we need to study these a bit closer. It turns out that the metric perturbations depend, so to speak, on one's point of view, so in order to make any sense of the equations that we will find, we must present the usual way of classifying the various point-of-views.

3.2 Choosing a gauge. Metric perturbations

In a homogeneous and isotropic universe such as the FRW universe, there is a preferred frame of reference: The frame that follows the motion of free-falling observers measuring zero momentum density. In a FRW universe, such a frame is always possible to define (Liddle & Lyth, 2000).

However, when we do perturbation theory, it no longer becomes possible to find such a preferred frame of reference, since the universe is no longer strictly homogeneous nor isotropic, and it is then impossible to measure zero momentum density everywhere. This means we have some freedom to choose our coordinate system, and a specific choice of coordinates is called a *gauge* in this context.

When we earlier stated that perturbation theory involves perturbing all relevant quantities, this also entails the metric tensor. However, the specific form of the metric perturbation depends on the gauge in which we wish to work. Leaving the choice of gauge open for the moment, we can write down the most general form of the metric perturbation (Dodelson, 2003): $g_{\mu\nu} \rightarrow g^{(0)}_{\mu\nu} + \delta g_{\mu\nu}$, where $g^{(0)}$ is the FRW metric tensor as given in section 2.1.1, and δg is the first-order deviation from the FRW metric. In the next section, we shall look at the various forms this perturbation can take, and the relations between them.

3.3 Scalar, vector and tensor perturbations

It is convenient to classify the perturbations to the metric in three categories: *scalar, vector,* and *tensor* perturbations[1] (Mukhanov et al., 1992).

3.3.1 Scalar perturbations

The scalar perturbations are classified by four scalar functions ϕ, ψ, B, and E, and their derivatives. The most general form of the scalar perturbations can be written as (Mukhanov et al., 1992)

$$\delta g_{\mu\nu} = \begin{pmatrix} -2\phi & -aB_{,i} \\ -aB_{,i} & 2a^2(\psi\delta_{ij} - E_{,i,j}) \end{pmatrix}. \tag{3.1}$$

3.3.2 Vector perturbations

Vector perturbations are classified by two three-vectors S_i and F_i, both of which must satisfy $S_i^{,i} = F_i^{,i} = 0$. That is, they must be divergenceless; otherwise they could have been derived from a scalar function. The most general vector perturbation is

$$\delta g^V_{\mu\nu} = -\begin{pmatrix} 0 & -aS_i \\ -aS_i & a^2(F_{i,j} + F_{j,i}) \end{pmatrix}. \tag{3.2}$$

3.3.3 Tensor perturbations

Tensor perturbations are only classified by one tensor function h which must be symmetric and satisfy $h^i_i = 0, h^{,j}_{ij} = 0$ (traceless and divergenceless) if it is to be a pure tensor perturbation. We then write the general tensor perturbation as

$$\delta g^T_{\mu\nu} = -a^2 \begin{pmatrix} 0 & 0 \\ 0 & h_{ij} \end{pmatrix}. \tag{3.3}$$

3.3.4 The decomposition theorem

One of the reasons this kind of decomposition is useful, is that when we plug the metric, perturbed as described above, into the Einstein equations, and

[1] The names refer to how the perturbations transform under a coordinate transformation on a *constant-time* hypersurface (Mukhanov et al., 1992).

perturb the energy-momentum tensor, a similar decomposition of it (the energy-momentum tensor) can be made. It then turns out (Liddle & Lyth, 2000) that the scalar, vector, and tensor perturbations evolve independently of each other. This means that we can find the equations that govern, say, scalar perturbations, without bothering to worry whether the tensor or vector perturbations will have any effect - they won't! And if there is any of the components that we have no interest in calculating, we won't need to.

Incidentally, there is one component that won't interest us in this text: vector perturbations. We will ultimately be concerned with the ratio of tensor-to-scalar-perturbations (whatever that might mean at present), and we are implicitly considering only Universe models in which no vector perturbations arise (to first order). Thus, from here on out, we won't mention vector perturbations again.

3.4 Gauge transformations

We now need some prescriptions for how to relate quantities defined in one gauge in terms of quantities defined in another - we must know how to do a *gauge transformation*. Let (t, x^i) denote our original coordinate system, and $(\tilde{t}, \tilde{x}^i)$ the coordinates in the coordinate system to which we wish to transform. If the transformation can be described by a small perturbation (which is appropriate when we are doing perturbation theory), that is, if

$$\begin{aligned} t &\rightarrow \tilde{t} = t + \xi^0(t, \vec{x}) \\ x^i &\rightarrow \tilde{x}^i = x^i + \delta^{ij}\xi_{,j}(t, \vec{x}), \end{aligned}$$

then the scalar perturbations transform as (Dodelson, 2003)

$$\begin{aligned} \phi &\rightarrow \tilde{\phi} = \phi - \frac{1}{a}\dot{\xi}^0 \\ B &\rightarrow \tilde{B} = B - \frac{\xi}{a} + \dot{\xi} \\ \psi &\rightarrow \tilde{\psi} = \psi - H\xi^0 \\ E &\rightarrow \tilde{E} = E + \xi. \end{aligned} \tag{3.4}$$

We also need to know how the tensor perturbations transform under a general gauge transformation. It is shown in Liddle & Lyth (2000) that the tensor perturbations are in fact gauge invariant, so the answer is that they don't at all. Knowing this, we will write down the tensor perturbations again, this time in a form that we will use in what follows, and which will

prove convenient:

$$h_{ij} = a^2 \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.5}$$

This choice of h_{ij} corresponds to a choice of coordinate axes: When we Fourier transform the h_ϵ's ($\epsilon = +, \times$), we choose our axes so that the wave vector always points in the $\hat{z}$ direction.

Gravitational waves

If we insert the above form of the tensor perturbations into the Einstein equations, it is shown in Dodelson (2003) that these perturbations, in Fourier space, must satisfy

$$\ddot{h}_\epsilon + 2\frac{\dot{a}}{a}\dot{h}_\epsilon + k^2 h_\epsilon = 0, \quad \epsilon = +, \times. \tag{3.6}$$

This is a wave equation with a damping term, and for that reason, the tensor perturbations to the metric are also called *gravitational waves*. We will study this equation in more detail in Part II.

Returning to scalar perturbations, we will now mention two gauges which are commonly used, and which will be relevant to this text. We will also mention a formalism which defines certain variables that will be independent of gauge choice.

3.4.1 The conformal Newtonian gauge

This gauge, which is also called the longitudinal gauge, is characterised by letting $B = E = 0$, and commonly used notation is to put $\phi = \Psi$, $\psi = \Phi$. The total (unperturbed + perturbed) line element then looks like

$$ds^2 = -(1 + 2\Psi)dt^2 + a^2(1 + 2\Phi)\delta^{ij}dx_i dx_j. \tag{3.7}$$

The perturbation Ψ is the generalization of the Newtonian gravitational potential, and Φ is a perturbation to the spatial curvature. The Newtonian gauge is mathematically the most simple gauge, and is what we will use in most of this text.

3.4.2 The synchronous gauge

We should really be calling this the synchronous *gauges*, because there are indeed an infinite amount of gauges which satisfy the synchronous gauge conditions (Liddle & Lyth, 2000), which are $\phi = B = 0$. This gauge was used by the whole cosmological community before the paper Bardeen (1980) came out, which gave a gauge-invariant description of perturbation theory. This gauge is more effective than the Newtonian gauge for numerical computations (Liddle & Lyth, 2000), and programs such as CAMB (Lewis et al., 2000) use it. These gauges can be seen as constant-time hypersurfaces in space-time, so that wherever we are situated on such a hypersurface, the time coordinate will always be the same. The multiple synchronous gauges reflect the fact that we can choose any time to be that constant time coordinate.

3.4.3 Gauge-invariant formalism

Using eq. (3.4), it is possible to construct certain gauge invariant variables from the scalar perturbations, that is, variables that do not transform under a small change of coordinates (Bardeen, 1980):

$$\Phi_A \equiv \phi + \frac{1}{a}\frac{\partial}{\partial\eta}(a(\dot{E} - B)) \tag{3.8}$$

$$\Phi_H \equiv -\psi + aH(B - \dot{E}). \tag{3.9}$$

In the conformal Newtonian gauge, we see that $\Phi_A = \Psi$ and $\Phi_H = -\Phi$, which makes this gauge especially convenient to work with. Inserting the potential Φ_A into the Einstein equations, and limiting oneself to universes filled with one type of perfect fluid, one may derive the following evolution equation (Mukhanov et al., 1992):

$$\ddot{\Phi}_A + 3\alpha\left(1 + \frac{\dot{p}}{\dot{\rho}}\right)\dot{\Phi}_A + \frac{\dot{p}}{\dot{\rho}}k^2\Phi_A + \left(2\dot{\alpha} + \left(1 + 3\frac{\dot{p}}{\dot{\rho}}\right)\alpha^2\right)\Phi_A = 0. \tag{3.10}$$

The quantity α is defined as

$$\alpha \equiv \frac{\dot{a}}{a} = aH, \tag{3.11}$$

and is therefore the inverse comoving Hubble radius. We can also consider gauge invariant combinations of the energy-momentum tensor T, which, as we mentioned earlier, will also transform under a small coordinate transformation. The gauge-invariant velocity and energy density are defined as (Bardeen, 1980)

$$v \equiv ikB + \frac{\hat{k}^i T^0_i}{(\rho + \mathcal{P})a} \tag{3.12}$$

$$\epsilon_m \equiv -1 - \frac{T^0_{\ 0}}{\rho} + \frac{3H}{k^2\rho} k^i T^0_{\ i}. \tag{3.13}$$

This concludes this chapter. In the next three chapters, we will put the definitions and formalisms introduced in this chapter to use.

Chapter 4

Inflation

According to widely held beliefs in the scientific community, in the very early Universe there was a period of accelerated expansion, a period called *inflation*. Actually, inflation refers, originally, not only to that period, but to *any* period of accelerated expansion. That is, when $\frac{d^2a}{dt^2} > 0$, we have inflation (Liddle & Lyth, 2000).

The main merit of inflation in our context is its ability to provide a plausible mechanism for setting up the initial perturbations which are then subsequently governed by the equations of chapter 5. In this chapter we will take a closer look at this mechanism, and we will examine the simplest model of inflation and some of its characteristics in order to make clear what the fuss in Part II is all about.

4.1 Initial structure

The inflation model is capable of providing the initial perturbations that evolve into the structure we see today, by amplifying the quantum fluctuations that are present in any system, including the early Universe. In this section, we show how this comes about. First, however, we need to clarify some details:

First: As mentioned in the previous chapter, though the photon perturbations are our main point of interest, these depend on the other types of perturbations that might be present. In order to make any predictions for the photon perturbations today, then, we should need to know the initial conditions for perturbations of all types and flavors (like baryons, neutrinos,

dark matter, metric perturbations etc.). However, if one is willing to make certain assumptions and approximations[1], it is shown in Dodelson (2003) that once the initial conditions for *one* of the components are known, then all of the others are determined[2]. We will define a curvature perturbation ζ, in terms of which all other perturbations can be expressed, and examine the inflationary predications for it.

Second: We are looking at the initial conditions for modes with *large* wavelengths - that is, wavelengths that are larger than the horizon. As we noted in the beginning of the chapter, inflation is defined as any period where $d^2a/dt^2 > 0$, which translates into

$$\frac{d(H^{-1}/a)}{dt} < 0, \tag{4.1}$$

meaning that the comoving Hubble horizon actually *decreases* during inflation. This means that, during inflation, the Fourier modes corresponding to successively shorter wavelengths are pushed beyond the horizon, effectively 'freezing out' these wavelengths. We assume that inflation goes on long enough for all modes of interest to be well beyond the horizon at the end of inflation. As the radiation era sets in, the horizon starts increasing, so that larger and larger modes again may physically interact. We are interested in the initial conditions starting at the period betweeen horizon exit (during inflation) and horizon reentry - a period referred to as the *primordial epoch*. Since the modes outside the horizon cannot change due to causal physics, the conditions of the primordial epoch are the same as the conditions just before the mode exited the horizon during inflation. As we shall see below, these conditions are described by quantum mechanics.

There are many models of inflation, but we will restrict ourselves to one of the simplest: the single-scalar field model, the reason being that this model highlights all the aspects of inflation which are important for this work.

4.1.1 The generation of inflation by a scalar field

The general theory of fields is briefly reviewed in appendix B.6, and here, we will study a particular type of field - namely, a *scalar field*, which represents a spin-zero particle. We denote this field by ϕ. The Lagrangian density for such a field takes the form

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi), \tag{4.2}$$

[1] Specifically, early times, tight coupling of photons and baryons, radiation domination and adiabatic initial conditions.

[2] At least for scalar perturbations. Tensor perturbations need only one initial condition, regardless.

where V is some potential, also dependent on the types of particles at which we are looking. There are currently no consensus as to what the potential should be, so we will just leave it as V, as knowledge of its specific shape is not crucial for our purposes. From the above Lagrangian density, we use eq. (B.19) to find the energy-momentum tensor of the scalar field. From it, we then arrive at the following expressions for the energy density and pressure of the scalar field, assuming that the scalar field is homogeneous:

$$\rho_\phi = \frac{1}{2}(\phi')^2 + V(\phi) \tag{4.3}$$

$$p_\phi = \frac{1}{2}(\phi')^2 - V(\phi), \tag{4.4}$$

The first term on each right hand side is the equivalent of the kinetic energy of the scalar field. We noted in chapter 2 that in order for inflation (accelerated expansion) to occur, we need the Universe to be dominated by a substance whose equation of state satisfies $w < -1/3$. If the scalar field is that substance, we see that we get inflation if $\rho_\phi \approx -p_\phi$, that is, if the kinetic energy term in eqs. (4.3) and (4.4) is negligible compared to the potential energy term. This happens if the scalar field is slowly rolling down the potential - hence, this approximation is called the *slow-roll* approximation[3].

It is conventional to introduce two slow-roll parameters ϵ and δ:

$$\epsilon \equiv \frac{M_{\rm Pl}^2}{2}\left(\frac{1}{V}\frac{dV}{d\phi}\right)^2 \tag{4.5}$$

$$\delta \equiv M_{\rm Pl}^2 \frac{1}{V}\frac{d^2V}{d\phi^2}, \tag{4.6}$$

where the *reduced Planck mass* $M_{\rm Pl} = (8\pi G)^{-1/2}$, and it can be shown (Liddle, 1999) that the conditions $\epsilon \ll 1$ and $|\delta| \ll 1$, called the slow-roll conditions, are sufficient (but not necessary) for inflation to occur. Some other useful formulas derived using the slow-roll-approximation are

$$3H\phi' \simeq -\frac{dV}{d\phi} \tag{4.7}$$

and

$$\frac{d(\ln H)}{H dt} \simeq -\epsilon, \quad \frac{d(\ln \epsilon)}{H dt} \simeq 4\epsilon - 2\delta. \tag{4.8}$$

[3]The field cannot be stationary, for reasons explained in Dodelson (2003).

4.1.2 Dynamics of the scalar field

We introduce the perturbed scalar field as

$$\phi(\eta, \vec{x}) = \phi^{(0)}(\eta) + \delta\phi(\eta, \vec{x}). \tag{4.9}$$

The quantity $\phi^{(0)}$ is the homogeneous part driving inflation as discussed above. When working in the longitudinal gauge, as we are, the perturbation $\delta\phi$ will be a gauge-invariant quantity (Mukhanov et al., 1992). The scalar field will couple to the scalar metric perturbations, and in order to determine this coupling, we must insert everything into the Einstein equations.

Inserting the perturbed form of the scalar field into the Lagrangian, and again finding the energy-momentum tensor as done above, we find that energy-momentum tensor will then also consist of a homogeneous part and a perturbed part. The homogeneous components give nothing new (those components are what we used to find eqs. (4.3) and (4.4)), but the perturbed part must be equated with the perturbed part of the Einstein tensor, which will, in general, contain the metric perturbations in some way or other. This way, we can find relations between the scalar field and the metric perturbations. We won't write down the perturbed energy-momentum tensor for the scalar field, since it is not very interesting in itself. However, we *will* write down the two relevant equations we get from the Einstein equations (or, equivalently, the Euler-Lagrange equations). The first governs the evolution of the perturbed scalar field:

$$\ddot{\delta\phi} + 2\alpha\dot{\delta\phi} + k^2\delta\phi = 0. \tag{4.10}$$

The second governs the evolution of the scalar metric perturbations Ψ, and is just eq. (3.10) applied to the inflationary stage:

$$\ddot{\Psi} + 2\frac{\dot{\phi}^{(0)}}{a}\frac{\partial}{\partial\eta}\left(\frac{a}{\dot{\phi}^{(0)}}\right)\dot{\Psi} + k^2\Psi + 2\dot{\phi}^{(0)}\frac{\partial}{\partial\eta}\left(\frac{\alpha}{\dot{\phi}^{(0)}}\right)\Psi = 0. \tag{4.11}$$

In both of the above equations, note that we have used α, defined in eq. (3.11). It is, however, important to note that the above equation only holds for non-de Sitter universes - for a pure de Sitter space, $\Psi = 0$ and there are no perturbations to the metric.

Actually, eq. (4.10) is an approximation, or rather, two approximations: There appears another term on the left hand side of it, involving the double derivatives of the scalar field potential with respect to the scalar field, and there appear some terms involving the scalar metric perturbations. We neglect these because *if* we assume the field is light, or near massless, the potential term will be small, while the scalar metric terms will turn out to be negligible compared to the other terms.

The conservation of ζ

Using the above equations, we can arrive at a useful result: Define

$$\zeta \equiv \frac{2}{3}\frac{H^{-1}\Psi' + \Psi}{1+w} + \Psi. \tag{4.12}$$

Here, w is the coefficient in the equation of state, eq. (2.8). This quantity, called the *curvature perturbation*, is gauge-invariant. Furthermore, it satisfies

$$\frac{3}{2}\zeta' H(1+w) = -\frac{k^2}{a^2}\Psi. \tag{4.13}$$

For long-wavelength perturbations, where $k \to 0$, we see that ζ is a constant, and is thus conserved in time for such wavelengths. Finally, all of the relevant primordial perturbations may be expressed in terms of ζ, so that if we can find inflationary predictions for this quantity, we have found all the initial conditions we need.

4.1.3 The quantum mechanical generation of perturbations

Although we now know well how the perturbations to the metric and the scalar field evolve during inflation, we still have provided no mechanisms for making the perturbations appear in the first place. That is what we will do now, using quantum mechanics. We will, in this subsection, follow the treatment in Lyth & Liddle (2009), along with the definitions in Baumann et al. (2009).

We will now promote the scalar field to operator status, as described in appendix B.6, and see how this can generate initial perturbations. Below, we will show why we can use a well-known quantisation scheme, namely that of a harmonic oscillator, for the current problem.

Scalar perturbations

The equation we work with here is eq. (4.10). Introducing

$$\varphi(\eta) \equiv a\delta\phi(\eta), \tag{4.14}$$

the evolution equation for $\delta\phi$ then turns into

$$\ddot{\varphi} + \varphi\left(k^2 - \frac{\ddot{a}}{a}\right) = 0, \tag{4.15}$$

which can be seen as a harmonic oscillator equation with a time dependent frequency term. For scalar fields, it is, equivalently, an equation describing a scalar field with a time-dependent mass.

If we decide to work with Fourier modes much smaller than the horizon (this is equivalent to saying that all the scales relevant today were, at one point during inflation, much smaller than the horizon. In the inflationary model we are studying, this is a rather conservative assumption), it can be shown that $k^2 \gg \ddot{a}/a$, so we may approximate eq. (4.15) by a harmonic oscillator equation with a time-independent term, or, again, a scalar field with constant mass.

This motivates a quantisation procedure similar to that of a harmonic oscillator: We turn φ into an operator and write it as

$$\hat{\varphi}_{\vec{k}}(\eta) = \varphi_k(\eta)\hat{a}(\vec{k}) + \varphi_k^*(\eta)\hat{a}^\dagger(-\vec{k}), \tag{4.16}$$

where we have now made explicit the fact that we are looking at Fourier modes. The creation and annihilation operators satisfy standard commutation relations, and operators for different Fourier modes commute. The operator coefficients φ_k satisfy eq. (4.15). In the limit mentioned above, this is an equation for a constant-mass scalar field. There are two solutions to this equation, and standard procedure is to choose the solution

$$\varphi_k(\eta) = \frac{1}{\sqrt{2k}} e^{-ik\eta}, \tag{4.17}$$

with the normalisation constant chosen to make right the physical interpretation of the scalar field. With this initial value of the field, the more general solution to eq. (4.15) becomes

$$\varphi_k(\eta) = \frac{e^{-ik\eta}}{\sqrt{2k}} \frac{(k\eta - i)}{k\eta}. \tag{4.18}$$

As we mentioned in the beginning of the section, we are interested in the initial conditions valid at the primordial epoch, meaning the epoch where all relevant scales are beyond the horizon. This corresponds to taking the limit $k|\eta| \to 0$, which gives

$$\lim_{k|\eta|\to 0} \varphi_k(\eta) = \frac{e^{-ik\eta}}{\sqrt{2k}} \frac{-i}{k\eta}. \tag{4.19}$$

Tensor perturbations

The governing equation for tensor perturbations is eq. (3.6). Looking at eq. (4.10) and eq. (3.11), we see that the evolution equations for the scalar field and tensor perturbations are identical, so everything we said about scalar fields above is true also for tensor perturbations.

4.1.4 Power spectra and inflationary predictions

The above subsection treated the quantum-mechanical equations governing the perturbations during the inflationary epoch; now, we must translate these equations into observable predictions - namely, what can we predict for the values of the perturbations at the primordial epoch?

To answer this, we must first remember that quantum-mechanical operators are not observables - rather, their expectation values and variances are. In order to define these quantities, we must first specify the quantum-mechanical states in which to compute them. We work with states which are created and annihilated by the operators defined in eq. (4.16), and the corresponding equation for tensor perturbations. The ground (vacuum) state is the one satisfying $\hat{a}|0\rangle = 0$, and we assume that for all Fourier modes, we are in the ground state. The equation for $\hat{\varphi}$ is the same as that for an harmonic oscillator, meaning it is drawn from a Gaussian distribution centered at zero. The expectation value of the operator $\hat{\varphi}$, then, is zero, and we define its *power spectrum* $\mathcal{P}_\varphi$ by

$$\langle \hat{\varphi}_{\vec{k}} \hat{\varphi}_{\vec{k}'} \rangle = \frac{2\pi^2}{k^3} \mathcal{P}_\varphi(k) \delta^3(\vec{k} - \vec{k}'). \tag{4.20}$$

This again reflects the fact that we are working with a Gaussian distribution, and no correlation between the various $\vec{k}$ modes is assumed. Inserting the expression (4.16) for $\hat{\varphi}$, we find

$$\mathcal{P}_\varphi(k, \eta) = \frac{k^3}{2\pi^2} |\varphi_k(\eta)|^2, \tag{4.21}$$

which, when evaluated when eq. (4.19) is applicable, becomes

$$\mathcal{P}_\varphi(k) = \left(\frac{1}{2\pi\eta} \right)^2 = \left(\frac{a_k H_k}{2\pi} \right)^2. \tag{4.22}$$

The last equation follows from the fact that during inflation, the Hubble parameter H is approximately a constant, from which it follows that $\eta = -1/aH$. The subscript k means 'evaluated at horizon crossing'. Finally, we relate the power spectrum of φ to that of the original perturbation, $\delta\phi$, by dividing both sides by a^2:

$$\mathcal{P}_{\delta\phi}(k) = \left(\frac{H_k}{2\pi} \right)^2. \tag{4.23}$$

For tensor perturbations $h_{+,\times}$, we had exactly the same evolution equation as for $\delta\phi$, and thus end up with the same result, albeit with a different prefactor which can be derived using more advanced methods than those used here (Lyth & Liddle, 2009):

$$\mathcal{P}_h(k) = \frac{8}{M_{Pl}^2} \left(\frac{H_k}{2\pi} \right)^2. \tag{4.24}$$

The power spectrum for ζ

Since ζ is a conserved quantity in the long-wavelength limit (or, more relevant to the present discussion: in the primordial era), and all other relevant perturbations can be expressed in terms of it, it would be more helpful to have the power spectrum for ζ instead of that for $\delta\phi$. There are several ways of relating the two - here, we will still follow Lyth & Liddle (2009) with some input from Dodelson (2003).

Let us enter a gauge in which $\psi = E = 0$ - a *spatially flat* gauge. In this gauge, the evolution of $\delta\phi$ is still given by eq. (4.10). Further, ζ becomes (in terms of the scalar field)

$$\zeta = -aH\frac{\delta\phi}{\dot{\phi}^{(0)}}. \tag{4.25}$$

Thus, we can easily relate the spectrum of $\delta\phi$ in the spatially flat gauge (which is the same as the spectrum we derived above) to the spectrum of ζ:

$$\mathcal{P}_\zeta(k) = \left(\frac{a_k H_k}{\dot{\phi}_k^{(0)}}\right)^2 \mathcal{P}_{\delta\phi} = \frac{1}{4\pi^2}\left(\frac{a_k H_k^2}{\dot{\phi}_k^{(0)}}\right)^2. \tag{4.26}$$

We should, as indicated, evaluate both a, H, and $\dot{\phi}^{(0)}$ at horizon crossing. Thus, we have connected the spectrum of $\delta\phi$ to the spectrum of the conserved and gauge-invariant quantity ζ.

Predictions

It is conventional to parametrise the scalar and tensor power spectra as follows:

$$\mathcal{P}_\zeta(k) = A_s(k_p)\left(\frac{k}{k_p}\right)^{n_s(k_p)-1+\frac{1}{2}\alpha_s(k_p)\ln(k/k_p)}, \tag{4.27}$$

and

$$\mathcal{P}_h(k) = A_t(k_p)\left(\frac{k}{k_p}\right)^{n_t(k_p)}, \tag{4.28}$$

where k_p is the *pivot wavenumber*. These equations assume (based on empirical observation) that the spectra can be described by simple power-laws, characterised by the *spectral indices* n_s and n_t (these are also called *tilt*. An $n_s > 1$ is called a 'blue' spectrum, while an $n_s < 1$ is called a 'red' spectrum), plus the running of the spectral index α_s. These quantities are properly defined as

$$n_s - 1 \equiv \frac{d\ln\mathcal{P}_\zeta}{d\ln k}, \quad \alpha_s \equiv \frac{dn_s}{d\ln k}, \quad n_t \equiv \frac{d\ln\mathcal{P}_h}{d\ln k}. \tag{4.29}$$

We also note that since we assume both H and $\dot{\phi}^{(0)}$ to be approximately constant during inflation, we predict a nearly scale-invariant spectrum - one where $n_s \approx 1$ and $n_t \approx 0$ (the factor a in the expressions for the power spectra is absorbed into the conformal time derivative). Using the slow-roll equations (4.5) and (4.7), we can write eqs. (4.24) and (4.26) as

$$\mathcal{P}_\zeta = \frac{1}{24\pi^2 M_{Pl}^4}\frac{V_k}{\epsilon_k} \tag{4.30}$$

and

$$\mathcal{P}_h = \frac{2}{3\pi^2}\frac{V_k}{M_{Pl}^4}. \tag{4.31}$$

Calculating the spectral indices for these quantities, using eq. (4.8) and the fact that $d\ln(aH) \simeq H dt$ during inflation, we get

$$n_s(k) - 1 = -6\epsilon_k + 2\delta_k, \quad n_t(k) = -2\epsilon_k, \tag{4.32}$$

which provide the link between the observables (n_s and n_t) and the inflationary parameters. There is one more such link: the *tensor-to-scalar ratio* r, defined as

$$r \equiv \frac{\mathcal{P}_h}{\mathcal{P}_\zeta} = 16\epsilon_k = -8n_t. \tag{4.33}$$

The last equality ($r = -8n_t$) is called the *consistency relation* of inflation. It only holds for single-field inflation, which we have treated here; for inflation which arises from several fields, there is no such relation between the scalar and tensor fluctuations. A single-field inflationary model characterised by the slow-roll parameters or equivalent may thus be ruled out observationally, since there are three observables and two degrees of theoretical freedom. We observe that r is typically a small quantity as predicted by inflation.

Chapter 5

The evolution of photon perturbations

Now that we have provided a mechanism for setting up the initial perturbations in the universe, we need to find the equations that govern the evolution of the photon perturbations, so that we can see what they should look like today. In order to accomplish this, we turn to the Boltzmann equation (see appendix B.1).

5.1 Defining the problem

We first look at the scalar perturbations to the metric, and we will be interested in how they couple to the perturbations to the photon distribution, because these eventually become what we call the Cosmic Microwave Background, which is the only observable we will work with in this text. We must look at both the photon intensity and the polarisation, and provide the evolution equations for these.

Then, we look at the tensor perturbations and how they generate intensity and polarisation perturbations to the photon distribution. The set of equations that completely determine how the various photon distribution perturbations evolve are the goal of this chapter.

To zeroth order, the photon distribution obeys a Bose-Einstein distribution, which we will call $f^{(0)}(t)$. Note that it only depends on time, which is appropriate in a homogeneous and isotropic Universe. Clearly, in the unperturbed universe there is also no polarised light. We now move to Fourier

space and define intensity and polarisation perturbations by utilising the Stokes parameters (see appendix B.5): Let the first order perturbations to the polarisation density matrix be denoted $\delta\rho$, and the zeroth-order matrix $\rho^{(0)}$. The Stokes parameter brightness perturbations are then defined as (Kosowsky, 1996):

$$\begin{aligned}
\Theta^I(\vec{k},\vec{p},t) &= \left(pa\frac{\partial\rho_{11}^{(0)}(p,t)}{\partial(pa)}\right)^{-1}\Big(\delta\rho_{11}(\vec{k},\vec{p},t)+\delta\rho_{22}(\vec{k},\vec{p},t)\Big)\\
\Theta^Q(\vec{k},\vec{p},t) &= \left(pa\frac{\partial\rho_{11}^{(0)}(p,t)}{\partial(pa)}\right)^{-1}\Big(\delta\rho_{11}(\vec{k},\vec{p},t)-\delta\rho_{22}(\vec{k},\vec{p},t)\Big)\\
\Theta^U(\vec{k},\vec{p},t) &= \left(pa\frac{\partial\rho_{11}^{(0)}(p,t)}{\partial(pa)}\right)^{-1}\Big(\delta\rho_{12}(\vec{k},\vec{p},t)+\delta\rho_{21}(\vec{k},\vec{p},t)\Big) \quad (5.1)
\end{aligned}$$

Here, p (and $\vec{p}$) denotes the photon momentum, *not* some sort of pressure. We note that the brightness perturbations incorporate inhomogeneity (via the Fourier variable $\vec{k}$, conjugate to $\vec{x}$) and isotropy (via $\vec{p}$). For linear perturbations, the intensity brightness perturbation is equal to the fractional photon temperature perturbation $\delta T/T$.

The V (circular) polarisation is not expected to be present in the early Universe, so we only look at the I, Q and U parameters and their perturbations.

Finally, it is important to realise that the perturbation variables as defined in eq. (5.1) can be split up into scalar and tensor contributions, so that really, eq. (5.1) is the sum of the photon perturbations due to both scalar and tensor perturbations. This all follows from the decomposition theorem, described in chapter 3.

5.1.1 Boltzmann collision terms

In order to find the evolution equations for the photon field perturbations, the Boltzmann equation must be used. The relevant collision term to be inserted into eq. (B.1) will for photons be the Compton scattering terms. These are treated in e.g. Dodelson & Jubas (1995), and the only relevant information we need to note is that if we choose to work in a reference frame where the Fourier wave vector $\vec{k}$ is parallel with the z-axis, such collision terms only give rise to Q polarisation. Furthermore, the collisions are independent of the magnitude of the photon momentum, p. This again means that the perturbations defined in eq. (5.1) could be written with a $\hat{p}$-dependence instead of a $\vec{p}$-dependence, i.e., the perturbations only depend on the direction of photon momentum, not its magnitude. We will do this from

now on, but instead of writing this as $\Theta^s(\vec{k}, \hat{p}, \eta)$, we will write $\Theta^s(k, \mu, \eta)$, where $\mu = \hat{k} \cdot \hat{p}$.

5.2 Scalar perturbations

We start off with the scalar perturbations to the metric, and we will work in the conformal Newtonian gauge, as defined in chapter 3 (Eq. (3.7)).

5.2.1 Intensity

Setting up the Boltzmann equations for the intensity perturbations with Compton scattering terms taken into account is done in great detail in Dodelson (2003), and the resulting equation is

$$\dot{\Theta}^I + ik\mu\Theta^I = -\dot{\Phi} - ik\mu\Psi - \dot{\tau}(\Theta_0^I - \Theta^I + \mu v_b - \frac{1}{2}\mathcal{P}_2(\mu)\Pi). \tag{5.2}$$

Here, we are working in Fourier space, with conformal time η. τ is the *optical depth*, defined as $\tau = \int_\eta^{\eta_0} d\eta' n_e \sigma_T a$ where n_e is the electron number density and σ_T is the Thomson cross-section. v_b is the baryon velocity, which is typically small enough that we may treat it as a perturbation variable. Also in need of explanation is the quantity Θ_0^I, which is the *monopole* of the temperature perturbation. It forms the beginning of a sequence of *multipoles*, defined generally as

$$\Theta_l^s(k, \eta) \equiv \frac{1}{(-i)^l} \int_{-1}^{1} \frac{d\mu}{2} \mathcal{P}_l(\mu) \Theta^s(k, \mu, \eta), \tag{5.3}$$

where $s = I, Q, U$. The function $\mathcal{P}_l(\mu)$ here, and in the previous equation, is the l'th Legendre polynomial. Finally, we must define the quantity Π:

$$\Pi = \Theta_2^I + \Theta_2^Q + \Theta_0^Q, \tag{5.4}$$

providing the coupling between the photon intensity and polarisation, and reflecting our previous statement that Compton scattering only produces Q polarisation in the frame we have decided to work.

5.2.2 Polarisation

The only polarisation equation we get is an equation for the Q component of the perturbed polarisation field, and it is given in, e.g., Bond & Efstathiou (1984):

$$\dot{\Theta}^Q + ik\mu\Theta^Q = -\dot{\tau}(-\Theta^Q + \frac{1}{2}(1 - \mathcal{P}_2(\mu))\Pi). \tag{5.5}$$

5.3 Tensor perturbations

The tensor perturbations to the metric requires no specification of gauge, as we saw in chapter 3 that these perturbations are gauge-invariant. The perturbations are given by eq. (3.5), and in contrast to the scalar perturbations, the photon perturbations generated by tensor perturbations will depend not only on the angle between the photon direction and the z axis (here represented by μ), but also on the azimuthal angle, which we will call ϕ. We will denote the photon perturbations due to tensors by $\Theta^{s,T}, s = I, Q, U$ in order to more easily discern between these and the photon perturbations due to scalars.

It turns out that the evolution equations for the photon perturbations due to each of the two tensor modes separate (Crittenden et al., 1993), which we can take advantage of in the following way: Let us introduce new perturbation variables $\Theta_\epsilon^{s,T}, \epsilon = \{+, \times\}, s = \{I, Q, U\}$, defining them as follows:

$$\begin{aligned}
\Theta^{I,T} &= (1-\mu^2)\cos(2\phi)\Theta_+^{I,T} + (1-\mu^2)\sin(2\phi)\Theta_\times^{I,T} \\
\Theta^{Q,T} &= (1+\mu^2)\cos(2\phi)\Theta_+^{Q,T} + (1+\mu^2)\sin(2\phi)\Theta_\times^{Q,T} \\
\Theta^{U,T} &= -2\mu\sin(2\phi)\Theta_+^{U,T} + 2\mu\cos(2\phi)\Theta_\times^{U,T}.
\end{aligned} \tag{5.6}$$

The evolution equations for these variables then turn out to be (Crittenden et al., 1993):

$$\begin{aligned}
\dot{\Theta}_\epsilon^{I,T} &= -ik\mu\Theta_\epsilon^{I,T} - \frac{1}{2}\dot{h}_\epsilon + \dot{\tau}(\Theta_\epsilon^{I,T} - \Sigma_\epsilon) \\
\dot{\Theta}_\epsilon^{Q,T} &= -ik\mu\Theta_\epsilon^{Q,T} + \dot{\tau}(\Theta_\epsilon^{Q,T} + \Sigma_\epsilon) \\
\Theta_\epsilon^{U,T} &= \Theta_\epsilon^{Q,T} \\
\Sigma_\epsilon &\equiv \left(\frac{1}{10}(\Theta_\epsilon^{I,T})_0 + \frac{1}{7}(\Theta_\epsilon^{I,T})_2 + \frac{3}{70}(\Theta_\epsilon^{I,T})_4\right. \\
&\quad \left. -\frac{3}{5}(\Theta_\epsilon^{Q,T})_0 + \frac{6}{7}(\Theta_\epsilon^{Q,T})_2 - \frac{3}{70}(\Theta_\epsilon^{Q,T})_4\right).
\end{aligned} \tag{5.7}$$

These equations, together with eqs. (5.2) and (5.5), give the equations we need in order to evolve the photon perturbations from the time of horizon exit until today. In the next chapter, we will look at one approach to solving this problem.

Chapter 6

The CMB power spectrum

In chapter 5, we ended up with some evolution equations for both the intensity and polarisation photon distributions. We will further study these equations in this chapter, in order to connect the primordial power spectra to an important observable - the Cosmic Microwave Background (CMB)

The CMB is simply another word for the photons that pervade the Universe, most of which originate from the decoupling event mentioned in chapter 2. Thus, when we observe these photons, we are looking at relics from when the Universe was only 380,000 years old, and since the Universe was opaque before then, this is as far back in time as we can see. These photons, then, give us unique information about the early Universe.

The first indications of the CMB were reported in Penzias & Wilson (1965), and measurements of its spectrum was carried out by the FIRAS instrument on the COBE satelite, which established that the spectrum is as close to a blackbody spectrum as anything can get (Mather et al., 1994). This strongly supported the Hot Big Bang theory, as this theory predicts an early universe in full thermal equilibrium, giving the photons emitted at decoupling a blackbody spectrum.

The perturbations set up in the early Universe should translate into perturbations in the photon distribution, as we will see below. Thus, in order to gain some insight into those very early times, we need to measure the photon anisotropies and compare them to theory. We will for this purpose present the framework within which this is typically done, both for the intensity and the polarisation perturbations.

We must again say, as in chapter 3, that we will not present a complete treatment of how to calculate the photon distribution - for that, we would

need the complete evolution equations for the matter distribution and the metric perturbations as well. Our aim here is simply to show how the photon distribution can be a probe into the primordial power spectra, and to look at the distinct imprints the various primordial parameters make on the CMB through numerical calculations.

6.1 The spherical expansion of the CMB

6.1.1 Intensity

The photon intensity perturbation is a function of position, photon momentum direction, and time, $\Theta^I = \Theta^I(\vec{x}, \hat{p}, \eta)$. Here, Θ^I denotes the sum of contributions from scalar and tensor perturbations, $\Theta^I = \Theta^{I,S} + \Theta^{I,T}$, but we can always choose to only look at one of the contributions due to the decomposition theorem.

As explained in appendix B.5, the intensity Stokes parameter does not transform under rotations of the coordinate system tangent to the sphere - it is a scalar under such transformations. It can then be expanded in spherical harmonics on the sphere defined by $\hat{p}$ as follows:

$$\Theta^I(\vec{x}, \hat{p}, \eta) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{I,lm}(\vec{x}, \eta) Y_{lm}(\hat{p}). \tag{6.1}$$

As also noted, all of the information of the perturbations are contained in the $a_{I,lm}$'s, so they completely determine the intensity variations (to first order).

6.1.2 The polarisation field

We now turn to the perturbations to the polarisation of the CMB radiation field $\Theta^Q(\vec{x}, \hat{p}, \eta), \Theta^U(\vec{x}, \hat{p}, \eta)$, and to the problem of how to expand these on the sky sphere. We have from appendix B.5 that these Stokes parameters transform according to eq. (B.13) under a rotation by an angle ψ of the two-dimensional coordinate system in which they are defined. In our case, this system lies in the tangent plane of the point on the sphere determined by $\hat{p}$, and so, the combination $\Theta^Q \pm i\Theta^U$ is, by the definition in appendix B.4, a spin-2 quantity. As such, it may be expanded using the twice-spin-lowered-or-raised spherical harmonics ${}_{\pm 2}Y_{lm}$:

$$(\Theta^Q \pm i\Theta^U)(\vec{x}, \hat{p}, \eta) = \sum_{lm} a_{\pm 2,lm}(\vec{x}, \eta)({}_{\pm 2}Y_{lm}), \tag{6.2}$$

analogous to eq. (6.1) for the intensity case. Alternatively, one may act twice on $\Theta^Q \pm i\Theta^U$ with spin raising or lowering operators, and then expand the resulting function in regular spherical harmonics, as explained in appendix B.4:

$$\begin{aligned} \bar{\eth}^2(\Theta^Q + i\Theta^U)(\vec{x},\hat{p},\eta) &= \sum_{lm}\left(\frac{(l+2)!}{(l-2)}\right)^{1/2} a_{2,lm}(\vec{x},\eta)Y_{lm}(\hat{p}) \\ \eth^2(\Theta^Q - i\Theta^U)(\vec{x},\hat{p},\eta) &= \sum_{lm}\left(\frac{(l+2)!}{(l-2)}\right)^{1/2} a_{-2,lm}(\vec{x},\eta)Y_{lm}(\hat{p}), \end{aligned} \quad (6.3)$$

where we have expanded the resulting spin-zero functions in terms of the expansion coefficients for the functions with a finite spin, $a_{\pm 2,lm}$. Instead of working directly with these quantities, however, we will follow normal conventions (Zaldarriaga & Seljak, 1997) and define

$$\begin{aligned} a_{E,lm} &= -\frac{1}{2}(a_{2,lm} + a_{-2,lm}) \\ a_{B,lm} &= \frac{i}{2}(a_{2,lm} - a_{-2,lm}), \end{aligned} \quad (6.4)$$

where the letters E and B are chosen because the properties of these quantities under parity transformations are the same as those of electric (E) and magnetic (B) fields.

6.1.3 Properties of the expansion coefficients

In this section, we will mention some of the properties of the various expansion coefficients $a_{X,lm}, X = I, E, B$. We know from chapter 4 that the initial perturbations set up during inflation came from quantum fluctuations with a random Gaussian distribution. This sets certain constraints on the $a_{X,lm}$'s: Their mean value, averaged over the whole distribution of $a_{X,lm}$'s, is zero, while their variance satisfies (Dodelson, 2003)

$$\langle a_{X,lm} a^*_{X',l'm'}\rangle = \delta_{ll'}\delta_{mm'}C_l^{XX'}. \quad (6.5)$$

The quantity in brackets is called the power spectrum of the $a_{X,lm}$'s, and is one of the most important observables of the CMB. Implicit in this equation lies two assumptions: The variance is independent of m, and there is no covariance, neither between different m's nor l's. If $X \neq X'$, we have a *cross-correlation* power spectrum. From eq. (6.5) follows the relations

$$C_l^{XX'} = \frac{1}{2l+1}\sum_m \langle a_{X,lm} a_{X',lm}\rangle. \quad (6.6)$$

In order to determine the power spectrum, we can then measure the $a_{X,lm}$'s. There is an important point to stress in that context: The $a_{X,lm}$'s represent samples from a distribution with a certain variance C_l. As with any variance of a quantity, our estimated value of that variance will contain some uncertainty stemming from the fact that we only have access to a finite number of samples. This uncertainty is called *cosmic variance*, in an attempt to convey that our sample size is, and always will be, limited: There are only $2l+1$ $a_{X,lm}$'s to measure, and we can't move to another part of the Universe to measure the $a_{X,lm}$'s there since we would then have to move beyond our horizon. The concept of cosmic variance is often expressed mathematically as follows:

$$\frac{\Delta C_l}{C_l} = \sqrt{\frac{2}{2l+1}},$$

where ΔC_l is the deviation of the observed C_l ($|a_{X,lm}|^2$) from its true value. At low l we get the fewest samples, and the highest cosmic variance, as expressed in the equation.

6.2 Analytic expressions for the power spectra

Now that we have presented the framework of describing the radiation field of the CMB, we will now present the analytic expressions for the various C_l's. There are two major steps in this task: The actual solving of the equations given in chapter 5, and the relating of the variables in that chapter, and chapter 4, to the C_l's of this chapter. We will treat the scalar and tensor cases separately, since the scalar case is a bit more straightforward.

6.2.1 Scalar anisotropies

By formally integrating eqs. (5.2) and (5.5), we get (Zaldarriaga & Harari, 1995)

$$\Theta^I(k,\mu,\eta_0) = \int_0^{\eta_0} d\eta \tilde{S}(k,\mu,\eta) e^{ik\mu(\eta-\eta_0)-\tau(\eta)} \tag{6.7}$$

$$\Theta^Q(k,\mu,\eta_0) = \frac{3}{4}(1-\mu^2)\int_0^{\eta_0} d\eta e^{ik\mu(\eta-\eta_0)} g(\eta)\Pi(k,\eta), \tag{6.8}$$

where the source function $\tilde{S}(k,\mu,\eta_0)$ is defined as

$$\tilde{S}(k,\mu,\eta) = -\dot{\Phi} - ik\mu\Psi - \dot{\tau}\left(\Theta_0^I + \mu v_b - \frac{1}{2}\mathcal{P}_2(\mu)\Pi\right), \tag{6.9}$$

the *visibility function* $g \equiv -\dot{\tau}e^{-\tau}$, and Π was defined in eq. (5.4). It is possible, and preferrable, to transform the expression for the intensity perturbations as described in detail in Dodelson (2003): Since the integrand in the first integral above multiplies $e^{ik\mu(\eta-\eta_0)}$, we can integrate by parts, effectively replacing every occurrence of μ with $-\frac{1}{ik}\frac{d}{d\eta}$. The surface terms either vanish or have no angular dependence and thus cannot contribute to the angular power spectrum[1] so we end up with a new integral and a new source function:

$$\Theta^I(k,\mu,\eta_0) = \int_0^{\eta_0} d\eta e^{ik\mu(\eta-\eta_0)} S(k,\eta), \tag{6.10}$$

$$\begin{aligned} S(k,\eta) &= g(\eta)\left(\Theta_0^I + \Psi - \frac{i\dot{v}_b}{k} + \frac{1}{4}\Pi + \frac{3}{4k^2}\ddot{\Pi}\right) + e^{-\tau}(\dot{\Psi}-\dot{\Phi}) \\ &\quad +\dot{g}(\eta)\left(-\frac{iv_b}{k} + \frac{3}{4k^2}\dot{\Pi}\right) + \frac{3}{4k^2}\ddot{g}(\eta)\Pi. \end{aligned} \tag{6.11}$$

We now need to transform the variables for which we have solutions - eqs. (6.8) and (6.10) - into observables. We start by transforming the solutions from Fourier space to real space. We know that all the evolution equations in chapter 5 were calculated in the frame where $\vec{k}||\hat{z}$. For intensity perturbations, this is fine, because the intensity of the radiation field is rotationally invariant and thus independent of direction on the sky. However, the Q (and U) type polarisation is not rotationally independent, but rather transforms according to eq. (B.13). This means that for every Fourier mode, the polarisation perturbations as calculated in the $\vec{k}||\hat{z}$ frame must be rotated into a fixed frame. This will in general introduce a $\vec{k}$-dependent angle, and we will see below how to deal with this.

Following Zaldarriaga & Seljak (1997), we write

$$\begin{aligned} \Theta^I(\hat{p}) &= \int \frac{d^3\vec{k}}{(2\pi)^3}\zeta(\vec{k})\Theta^I(\eta=\eta_0,k,\mu) \\ (\Theta^Q \pm i\Theta^U)(\hat{p}) &= \int \frac{d^3\vec{k}}{(2\pi)^3}\zeta(\vec{k})e^{\mp 2i\psi}\Theta^Q(\eta=\eta_0,k,\mu). \end{aligned} \tag{6.12}$$

We must explain this equation in a bit more detail. First of all, we notice that the real space variables Θ^I and $\Theta^Q \pm i\Theta^U$ are now only functions of the photon direction $\hat{p}$ alone. This comes about because we are interested in what goes on at earth today, nailing down $\vec{x}=0$ and $\eta=\eta_0$. This explains why we have no factor of $e^{i\vec{k}\cdot\vec{x}}$ in our Fourier transform - it vanishes when we choose our position to be the origin.

[1] More precisely, they only contribute to the monopole of the spherical decomposition, which is unobservable, since it is just the deviation from the uniform blackbody temperature, averaged over all space.

The next feature to explain is the appearence of the $\zeta(\vec{k})$ factor. ζ here denotes, as in chapter 4, the curvature perturbation, and we normalise it so it satisfies

$$\langle\zeta(\vec{k})\zeta^*(\vec{k}')\rangle = (2\pi)^3\frac{2\pi^2}{k^3}\mathcal{P}_\zeta(k)\delta^3(\vec{k}-\vec{k}'), \tag{6.13}$$

where $\mathcal{P}_\zeta$ is the primordial power spectrum of the scalar perturbations, introduced in chapter 4. By including ζ in the definition of the real-space perturbations, we effectively extract the stochastic nature of the perturbations from the intensity and polarisation perturbations, so that the perturbations for which we have obtained solutions are independent of initial conditions[2]. This works since, as we said in chapter 4, all of the primoridal perturbations can be related to one another under certain approximations, so that the stochastic nature of one perturbation - say, ζ - are shared with all of the others.

At last, we comment briefly on the angle ψ appearing in the polarisation expressions. This is the k-dependent angle mentioned above, and it appears in the exponential as it does because we are forming the combinations $\Theta^Q \pm i\Theta^U$, which we know have that transformation property. However, the combinations $\Theta^Q \pm i\Theta^U$ are just Θ^Q in the coordinate frame mentioned above, which explains why only Θ^Q appears on the right hand side. The angle ψ will in general be quite difficult to calculate, so we will follow the approach of Zaldarriaga & Seljak (1997): If we were working with rotationally invariant perturbations (which we are, for the intensity case), the angle above would not appear. We could then calculate the power spectrum due to each Fourier mode first, and then integrate over all Fourier modes to produce the total power spectrum, since the power spectra (eq. (6.5)) are rotationally invariant, and because each Fourier mode is statistically independent (which follows from eq. (6.13) and the fact that all of the statistical properties of the perturbations are contained in ζ).

Luckily, we *can* form rotationally invariant quantities from functions like $\Theta^Q \pm i\Theta^U$, since they have a well-defined spin: we act on them with the spin raising and lowering operators, defined in eq. (B.6). Let us, then, focus on a single Fourier mode, $(\Theta^Q \pm i\Theta^U)(k,\mu)$ in the frame where $\vec{k}||\hat{z}$. In this frame, as we have seen in chapter 5, we only get Q-type polarisation from scalar perturbations, so that $(\Theta^Q \pm i\Theta^U) = \Theta^Q$. We also remember that Θ^Q is not a function of azimuthal angle ϕ, so from eq. (B.6), we get that $\bar{\eth}^2(\Theta^Q + i\Theta^U) = \eth^2(\Theta^Q - i\Theta^U)$, which again means (from eqs. (6.3) and (6.4)) that $a_{2,lm} = a_{-2,lm}$, and $a_{B,lm} = 0$ for these cases. That is, *scalar perturbations do not generate B type polarisation.*

[2]One might call this a redefinition of the perturbations we have worked with so far, so that, e.g. $\Theta^I \to \Theta^I_{new} = \Theta^I/\zeta$. This is the approach we will take for tensor perturbations.

We now go on to act twice on the polarisation perturbation with the spin raising operator (or the lowering operator, since they both give the same result), and get (Zaldarriaga & Seljak, 1997)

$$\Theta^E(k,\mu) = \frac{3}{4}\int_0^{\eta_0} d\eta g(\eta)\Pi(\eta,k)\left(1+\frac{\partial^2}{\partial(k\eta)^2}\right)^2((k^2(\eta-\eta_0)^2)e^{ik(\eta-\eta_0)}), \tag{6.14}$$

the superscript chosen in anticipation that the expansion coefficients for this perturbation will be the $a_{E,lm}$'s that we met earlier (give or take some normalisation factor). We now have two rotationally invariant quantities which we can calculate in the $\vec{k}||\hat{z}$ coordinate system. Expanding the intensity perturbation in spherical harmonics, calculating the power spectrum for each Fourier mode, and summing over all modes gives (Zaldarriaga & Seljak, 1997)

$$\begin{aligned} C_l^{II} &= \frac{1}{2l+1}\int\frac{d^3\vec{k}}{(2\pi)^3}\frac{2\pi^2}{k^3}\mathcal{P}_\zeta(k)\sum_m\left|\int d\Omega Y^*_{lm}(\hat{p})\int_0^{\eta_0} d\eta S(k,\eta)e^{ik\mu(\eta-\eta_0)}\right|^2 \\ &= 4\pi\int\frac{dk}{k}\mathcal{P}_\zeta(k)\left(\int_0^{\eta_0} d\eta S(k,\eta)j_l(k(\eta_0-\eta)),\right)^2 \end{aligned} \tag{6.15}$$

where $j_l(x)$ is the spherical Bessel function of order l, and the last equality follows from the relation

$$\int d\Omega Y^*_{lm}(\hat{p})e^{ix\mu} = \sqrt{4\pi(2l+1)}i^l j_l(x)\delta_{m0}, \tag{6.16}$$

which holds in the $\vec{k}||\hat{z}$ frame. In the last equality we have also switched the sign of the argument of the spherical Bessel functions in order to make easier the introduction of the radiation multipoles further down (this is allowed since $|j_l(x)| = |j_l(-x)|$). Doing the same (expanding, calculating power spectrum, and summing over modes) for the E perturbation defined in eq. (6.14) gives

$$C_l^{EE} = 4\pi\frac{(l+2)!}{(l-2)!}\int\frac{dk}{k}\mathcal{P}_\zeta(k)\left(\frac{3}{4}\int_0^{\eta_0} d\eta g(\eta)\Pi(\eta,k)\frac{j_l(k(\eta_0-\eta))}{(k(\eta_0-\eta))^2}\right)^2. \tag{6.17}$$

In order to generalise these results and be able to form cross correlations, we write

$$\begin{aligned} \Theta_l^I(k) &= \int_0^{\eta_0} d\eta S(k,\eta)j_l(k(\eta_0-\eta)) \\ \Theta_l^E(k) &= \sqrt{\frac{(l+2)!}{(l-2)!}}\int_0^{\eta_0} d\eta S_E(k,\eta)j_l(k(\eta_0-\eta)) \\ S_E(k,\eta) &= \frac{3g(\eta)\Pi(\eta,k)}{4(k(\eta_0-\eta))^2}. \end{aligned} \tag{6.18}$$

Θ_l^I are just the multipoles that we have used in our evolution equations in chapter 5, integrated over conformal time (see Dodelson (2003) for details on this). We then finally get

$$C_l^{XX'} = 4\pi \int \frac{dk}{k} \mathcal{P}_\zeta(k) \Theta_l^X(k) \Theta_l^{X'}(k), \tag{6.19}$$

where $X, X' = I, E$ (B modes do not contribute for scalar modes, as we remember).

6.2.2 Tensor anisotropies

We proceed along the same lines for tensor anisotropies as for scalar anisotropies, though there are some subtleties that we must note as we go.

First, we will again keep track of the *statistical* properties of the perturbations through the quantities $\xi_+, \xi_\times$, as in the previous section. We note from eqs. (5.7) that $\Theta_\epsilon^{Q,T} = \Theta_\epsilon^{U,T}$ for each of the polarisations $\epsilon = +, \times$, so defining $\Theta^{P,T}(\eta, \mu, k)$ through

$$\xi_\epsilon \Theta^{P,T} = \Theta_\epsilon^{s,T}, \quad s = Q, U, \quad \epsilon = +, \times, \tag{6.20}$$

and writing for the intensity perturbation

$$\xi_\epsilon \Theta^{I,T} = \Theta_\epsilon^{I,T}, \quad \epsilon = +, \times, \tag{6.21}$$

we collect the polarisation information in the quantities ξ_ϵ. This is permitted because both polarisations follow the same evolution, according to eq. (5.7), so the only thing left to discern the two polarisations is their statistical properties (in other words, their initial conditions).

We then write the intensity perturbation and the combination $\Theta^{Q,T} \pm i\Theta^{U,T}$ in the following way:

$$\begin{aligned} \Theta^{I,T}(\eta, \mu, k) &= ((1-\mu^2) e^{2i\phi} \xi^1(\vec{k}) + \\ &\quad (1-\mu^2) e^{-2i\phi} \xi^2(\vec{k})) \Theta^{I,T}(\eta, \mu, k) \\ (\Theta^{Q,T} \pm i\Theta^{U,T})(\eta, \mu, k) &= ((1 \mp \mu)^2 e^{2i\phi} \xi^1(\vec{k}) + \\ &\quad (1 \pm \mu)^2 e^{-2i\phi} \xi^2(\vec{k})) \Theta^{P,T}(\eta, \mu, k), \end{aligned} \tag{6.22}$$

where the quantities ξ^1 and ξ^2 are defined as

$$\xi^1 = \frac{\xi^+ - i\xi^\times}{\sqrt{2}}, \quad \xi^2 = \frac{\xi^+ + i\xi^\times}{\sqrt{2}}. \tag{6.23}$$

Their relation to the tensor power spectrum is

$$\begin{aligned}\langle\xi^1(\vec{k})\xi^{1*}(\vec{k}')\rangle = \langle\xi^2(\vec{k})\xi^{2*}(\vec{k}')\rangle &= (2\pi)^3\frac{2\pi^2}{k^3}\frac{\mathcal{P}_h(k)}{2}\delta^3(\vec{k}-\vec{k}') \\ \langle\xi^1(\vec{k})\xi^{2*}(\vec{k}')\rangle = \langle\xi^2(\vec{k})\xi^{1*}(\vec{k}')\rangle &= 0. \end{aligned} \tag{6.24}$$

Integrating eq. (5.7) in the same way as for the scalar case (meaning, using integration by parts and so on), and inserting the solutions into eq. (6.22), we get the present-day Fourier modes for tensor perturbations, just as in eqs. (6.8) and (6.10). We won't write these down, since they have the same structure as eq. (6.22).

As for the scalar case, we must now act twice on $\Theta^{Q,T} \pm i\Theta^{U,T}$ with the spin raising or lowering operators in order to make this quantity rotationally invariant. This time, as opposed to the scalar case, we get a B mode contribution. Finally, we can again calculate the power spectra, this time due to tensor perturbations, and the result is (Zaldarriaga & Seljak, 1997)

$$C_l^{XX',T} = 4\pi\int\frac{dk}{k}\mathcal{P}_h(k)\Theta_l^{X,T}\Theta_l^{X',T}, \quad X, X' = \{I, E, B\}, \tag{6.25}$$

where we have

$$\begin{aligned}\Theta_l^{I,T} &= \sqrt{\frac{(l+2)!}{(l-2)!}}\int_0^{\eta_0} d\eta S^{I,T}(k,\eta)\frac{j_l(k(\eta-\eta_0))}{(k(\eta-\eta_0))^2} \\ \Theta_l^{(E,B),T} &= \int_0^{\eta_0} d\eta S^{(E,B),T}(k,\eta) j_l(k(\eta-\eta_0)), \end{aligned} \tag{6.26}$$

and the source functions are defined as

$$\begin{aligned} S^{I,T}(k,\eta) &= -\frac{\dot{h}}{2}e^{-\eta} + g\Sigma \\ S^{E,T}(k,\eta) &= g\left(\Sigma - \frac{\ddot{\Sigma}}{k^2} + \frac{2\Sigma}{(k(\eta-\eta_0))^2} - \frac{\dot{\Sigma}}{k^2(\eta-\eta_0)}\right) \\ &\quad -\dot{g}\left(\frac{2\dot{\Sigma}}{k^2} + \frac{4\Sigma}{k^2(\eta-\eta_0)}\right) - 2\ddot{g}\frac{\Sigma}{k} \\ S^{B,T}(k,\eta) &= g\left(\frac{4\Sigma}{k(\eta-\eta_0)} + \frac{2\dot{\Sigma}}{k}\right) + 2\dot{g}\frac{\Sigma}{k}, \end{aligned} \tag{6.27}$$

where $h \equiv h_+ = h_\times$, since the evolution equations for each component are the same.

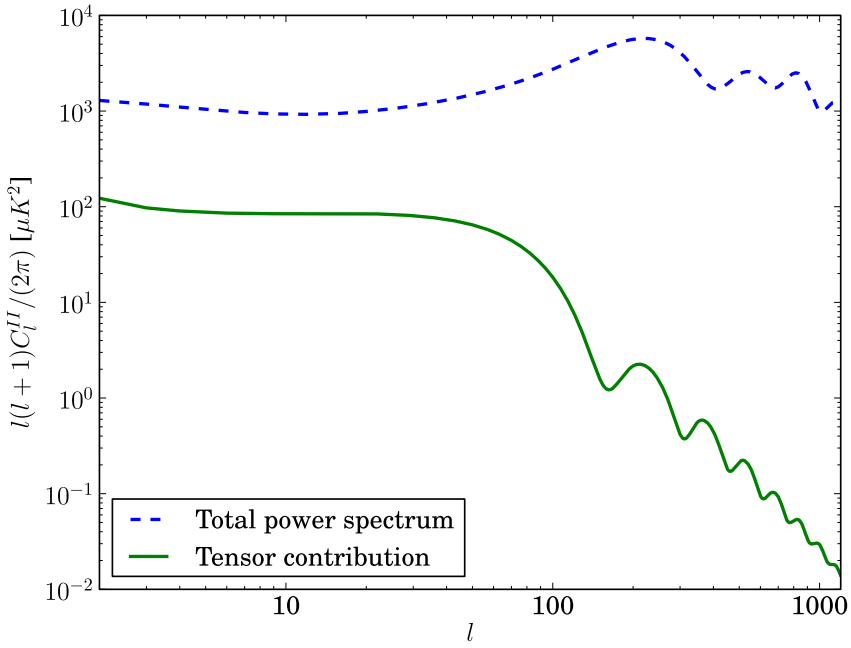

Figure 6.1: The total angular II power spectrum and the tensor contribution to that power spectrum for a 'typical' cosmological model with $r = 0.2$.

6.3 Detecting tensor contributions to the CMB

We have now seen how the primordial spectra for the curvature and tensor perturbations enter in the expressions for the present photon distribution. In this section, we will discuss the prospects of observing the tensor contribution to the angular power spectrum, using the knowledge that the primordial tensor-to-scalar ratio is typically predicted to be a small number. All of the results in this section is computed using *CAMB* software (Lewis et al., 2000).

6.3.1 Cosmological parameters

In order to make sense of the discussion in this section, we must first clarify the meaning of the most commonly used parameters that are input in programs such as *CAMB*.

We have, of course, the primordial parameters we have already met in chap-

ter 4: A_s, n_s, n_t, r (A_t is then determined by fixing r at the pivot wavenumber k_p). Further, we have some cosmological parameters, such as the various density parameters Ω and the present-day Hubble parameter H_0. In addition, one parameter in particular describes the reionisation event mentioned in chapter 2: the optical depth back to the time of reionisation, which henceforth is what is meant by τ.

There are many more parameters that influence the power spectrum, but the above are the ones we will mainly encounter in this text, so for now, we limit our discussion to these, and use best-fit values (Komatsu et al., 2010) for the rest.

6.3.2 The intensity spectrum

Looking at fig. 6.1, we see that for a typical model with a rather high value of the tensor-to-scalar ratio, the tensor contribution to the intensity power spectrum is predicted to be low, and its only contribution is to the range $l \lesssim 100$. This makes detecting tensor contributions to the intensity spectrum very difficult, for several reasons. First, at low l's, we have seen that the uncertainty due to cosmic variance is most pronounced, so that even if there is a tensor contribution at these l's, the uncertainty due to cosmic variance is due to be larger than the tensor contribution, and it is consequently impossible to tell whether we have a tensor contribution or not.

Second, there are several cosmological parameters that influence the low l range. As an illustration, fig. 6.2 shows that three distinct cosmological models produce spectra that are consistent with the error bars from the *WMAP* satelite data (NASA, 2010). By increasing τ, we are effectively increasing the rate at which the CMB interacts with the electrons in the Universe. This smoothes out anisotropies, reducing the power spectrum, but only on those scales that were within the horizon at reionisation, so only the higher l's are affected. Thus, by increasing the scalar amplitude, raising the whole spectrum, and then increasing τ, we can end up with a spectrum very similar to one where tensors are present. Raising n_s gives more power on smaller scales, as eq. (4.27) shows, which gives a similar (but not quite the same) effect as raising τ. In short, there are several parameters that can be tweaked to produce nearly the same power spectrum, so extracting the tensor contribution to the power spectrum can be very hard unless these other parameters can be fixed by other means.

To sum up: The II power spectrum seems to hold little promise in discovering the tensor contribution to the CMB, so we should look to the polarisation

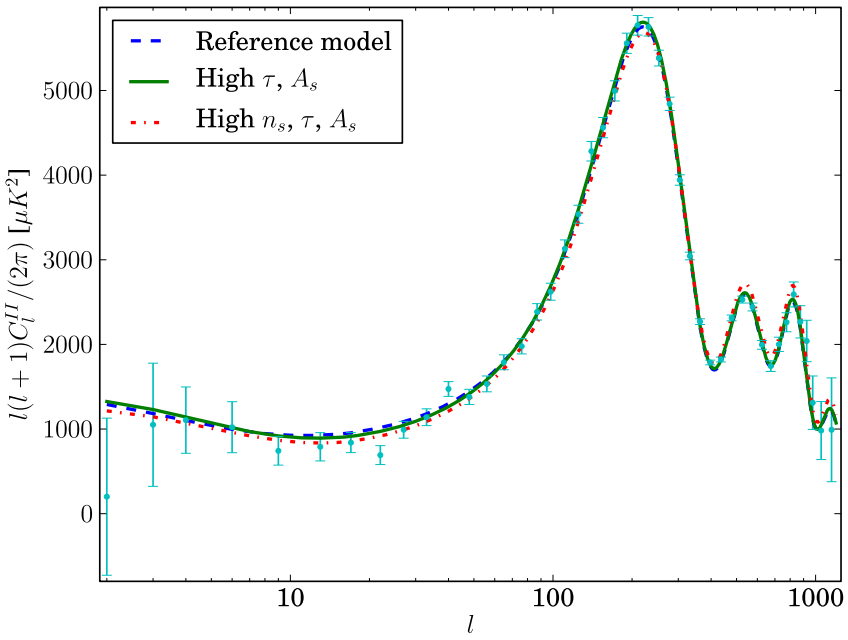

Figure 6.2: II power spectra of three cosmological models: The typical model that was shown in fig. 6.1, a model with high values of τ and the scalar amplitude A_s, and a model with high values of n_s, τ, and A_s. The two 'nontypical' models have no tensor contributions. Also shown is the binned error bars from the 7-year *WMAP* satelite data, including both cosmic variance and instrumental noise.

spectra and see what these may tell us.

6.3.3 The polarisation spectra

There are a number of spectra and cross-spectra one may form when including the polarisation spectra. The IB and EB spectra vanish for symmetry reasons (Baumann et al., 2009), so we will investigate the others here. As only high-l binned noise values for the EE and BB spectra have been made publicly available, we will not draw the noise bars for these spectra, but we will sum up their experimental status as we go along.

We start with the EE spectrum. We saw that both scalars and tensors contributed to this spectrum, and we see from fig. 6.3 that a high-r model leaves a relatively distinct imprint on this spectrum. However, the no-tensor

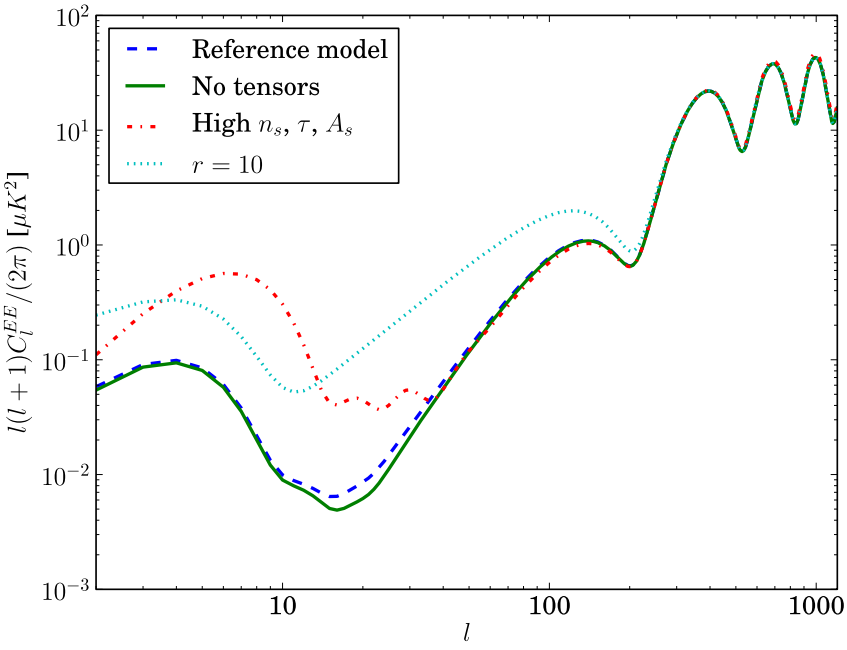

Figure 6.3: EE power spectra for the same reference models as in fig. 6.2, with an extra high tensor ($r = 10$) model.

model gives a spectrum that is virtually indistinguishable from a typical spectrum predicted by inflation, and we must go to unacceptably large r values in order to get a detection of tensors in this spectrum.

At any rate, the noise values for the EE spectrum are currently very large, so we must wait for more accurate experiments (e.g. *Planck* (Planck collaboration, 2005) and *QUIET* (Samtleben & QUIET collaboration, 2008)) in order to even be able to hint at some shape for this spectrum (Baumann et al., 2009).

For the IE spectrum, the experimental situation is different. As shown in fig. 6.4, the error bars are small enough that one might actually rule out some models - such as the reference model introduced earlier, with r now changed to 10. We see again that the contribution due to tensors changes the spectrum into a distinctive shape, as long as the primoridal tensor amplitude is high enough. Unfortunately, again we see that the differences between a typical model with no tensor contribution and a typical model with $r = 0.2$ are indistinguishable with the data we have.

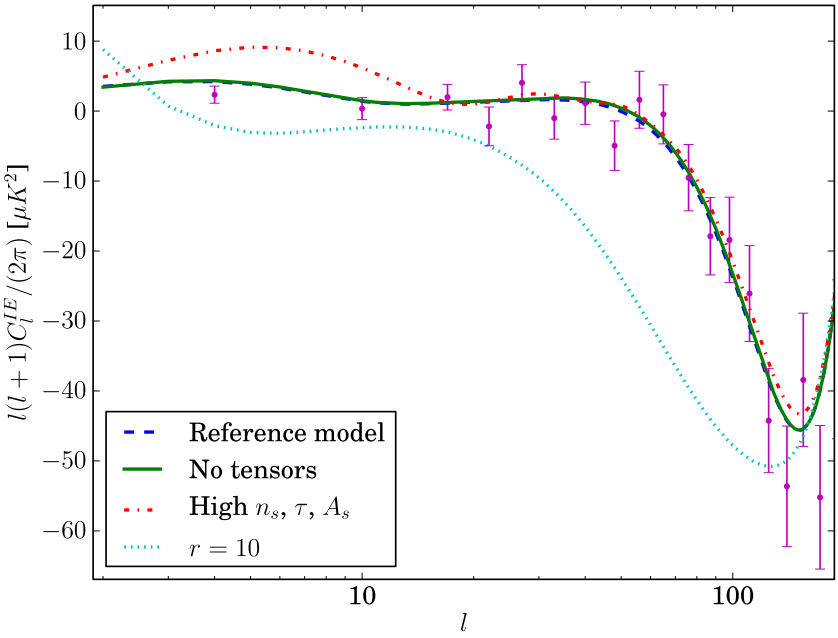

Figure 6.4: IE power spectra for the reference model and the high n_s, τ and A_s model from fig. 6.2, in addition to a no-tensor and $r = 10$ version of the reference model. Binned error bars from the 7-year *WMAP* experiment are shown, including both cosmic variance and experimental noise.

Finally, for the BB spectrum, we have a unique situation, in that, as we earlier found, only the tensors contribute to the B mode. Thus, if we have a detection of the BB power spectrum, we should be able to say that we have detected a tensor contribution to the CMB.

Unfortunately, there are complications also for this spectrum: There is another process which is able to induce B mode polarisation, called *gravitational lensing*, usually just 'lensing'. We will not do a thorough treatment of lensing here (see Lewis & Challinor (2006) for a review). What we need to note is that lensing essentially is the distortion of the CMB signal caused by massive objects between us and the *last scattering surface* - which is what we call the point from which the photons that constitute the CMB originated. The gravitational potentials of these massive objects cause a distortion of the light waves, effectively inducing B mode polarisation[3].

[3]The temperature and E mode (and their cross-spectra) are also affected, but the temperature spectrum is only marginally influenced until $l \approx 1000$ (Lewis & Challinor,

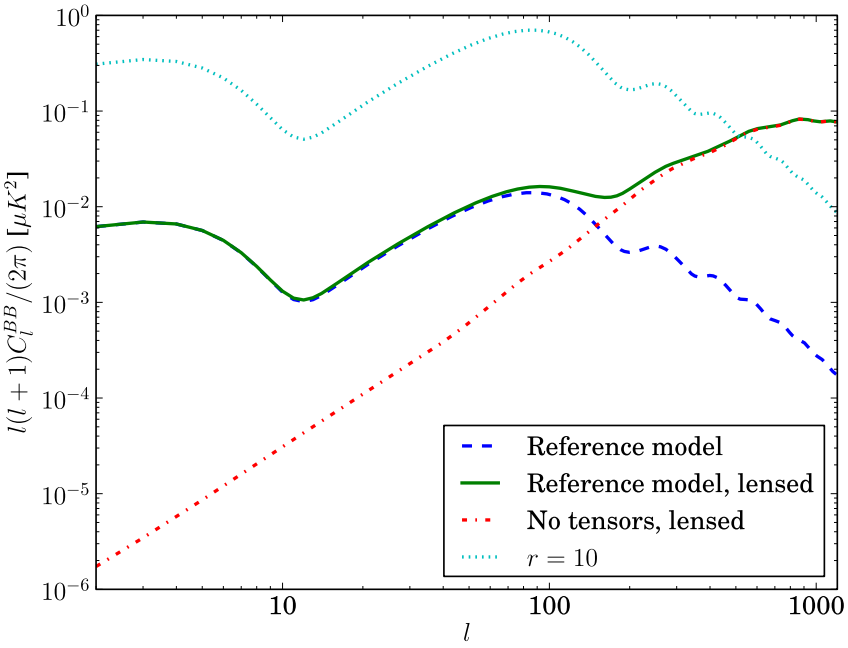

Figure 6.5: BB power spectra for the reference model from fig. 6.2, the lensed version of this model, a no-tensor lensed typical model, and a $r = 10$ typical model.

Fig. 6.5 shows that lensing is most pronounced at small scales, which makes sense intuitively - large-scale modes have not had the time to interact gravitationally with much since they entered the horizon. For a typical $r = 0.2$ model, lensing begins to dominate at $l = 100-200$, and above these l-values, even a no-tensor model can yield a nonzero BB spectrum.

For the moment, however, the limiting factor for BB detection is instrumental noise, not lensing contamination (Smith et al., 2008). If we manage to attain a sufficient degree of accuracy, we must use methods of 'delensing' in order to extract the primordial signal. This is accomplished through mapping of the structures that contribute to lensing, and is an ongoing effort.

2006), beyond which any primordial features have surely been washed out anyway, while the E mode is currently less interesting in that the uncertainties are still so large, and as we saw, it gives no unambigious sign of tensor contributions.

6.4 Chapter summary and current status

We have seen in this chapter that the present CMB angular power spectrum depends on the primordial spectra set up during inflation. With the primordial tensor perturbations in mind, we have studied the prospects of being able to detect this contribution to the CMB, and the difficulties associated with this task.

The current observational evidence, including the $WMAP$ observations in addition to other relevant observations (the baryonic acoustic oscillations (Percival et al., 2010), and supernova measurements (Hicken et al., 2009)) only gives an upper limit to the tensor-to-scalar ratio r evaluated at $k_p = 0.002$ Mpc^{-1}: $r < 0.20$. Simulations of the results from the ongoing $Planck$ mission (Planck collaboration, 2005) indicate that for a true tensor-to-scalar ratio of as small as 0.05, data obtained from $Planck$ should be able to put a lower limit on this, and improving the results of the $WMAP$ 5-year data by as much as a factor 9 (Colombo et al., 2009). Also, the $QUIET$ experiment (Samtleben & QUIET collaboration, 2008) which is devoted more exclusively to the polarisation of the CMB, will obtain better limits on the E and B modes than the ones we currently have.

Part II

Grishchuk's heresy

Chapter 7

Overview and Background

This Part is devoted to the presentation and examination of various controversial claims made by Leonid P. Grishchuk, especially the ones concerning the tensor-to-scalar ratio mentioned in chapter 4. This chapter will form part of the 'presentation' bit, and in it, we will present a short overview of Grishchuk's work up until the early 90's. Around that time, Grishchuk started expressing opinions that diverged from those of the cosmological community, summarised in Grishchuk (1994). We will give a short superficial overview of the controversy that followed, up until the status quo.

Since Grishchuk's claims take a stab at the heart, or at least a vital part, of the Standard Model of cosmology - the predicted relations between the primordial spectra - it is very important to sort out the truth of these claims. If they turn out to be true, a revision of the theories that are based on inflationary predictions about the primordial power spectra - quite a few, for instance theories of structure formation - are in order.

It is important to note that most of Grishchuk's academic work is uncontroversial, and that the main point of controversy is his claims concerning the tensor-to-scalar ratio, as we shall see below. He has, as we shall also see, derived important results, and there is every reason to take his claims seriously.

7.1 Adiabatic amplification of gravitational waves

The paper Grishchuk (1974) is our starting point, and it was one of the first papers by Grishchuk on the topic of gravitational waves, with which he has

later worked extensively. We will present the results and derivations found there, but we will present it using the notation of this text, and we will make use of the clarifications and elaborations found in Grishchuk (1993b).

We define μ^{gw} as

$$\mu^{gw}(\eta) \equiv a(\eta)h(\eta), \tag{7.1}$$

where h can be either of the two tensor polarisations, since they follow the same evolution equation. Eq. (3.6) becomes, in terms of this variable, and similar to eq. (4.15),

$$\ddot{\mu}^{gw} + \mu^{gw}\left(k^2 - U(\eta)\right) = 0, \tag{7.2}$$

where $U(\eta) \equiv \ddot{a}/a$. We see that for $k \gg |U(\eta)|$, this is a simple wave equation, and we find the result that the gravitational wave evolves adiabatically as $h \sim \exp(\pm ik\eta)/a$ - referred to by Grishchuk as the *adiabatical law* [sic]. However, for $k \ll |U(\eta)|$, there will be two solutions to this equation, corresponding to a growing and a decaying mode:

$$\mu_1^{gw} = a, \quad \mu_2^{gw} = a\int \frac{d\eta}{a^2}. \tag{7.3}$$

If we view the function $U(\eta)$ as a potential, we can think of a high-frequency wave as 'free', while a low-frequency wave might be interacting with the potential, entering the potential barrier at a time η_i, characterised by $k^2 = U(\eta_i)$, and exiting the barrier at a later time η_f. In the barrier region, the wave will consist of a linear combination of the solutions above, while outside it will decay adiabatically. The amplitude of the wave as it exits the barrier region depends on the entering conditions, but averaging the exiting amplitude over initial phases shows that the growing mode will give the dominant contribution. Thus, a typical μ^{gw} will go as a, and so the gravitational wave h will stay constant while the wave is below the barrier, meaning that it will be *superadiabatically amplified* (that is, amplified relative to a wave that obeys the adiabatical law) compared to a wave not interacting with the barrier at all. This amplification is driven by the potential $U(\eta)$, which is dependent only on the background metric. Thus, in a sense, the gravitational field provides the mechanism for amplifying gravitational waves.

7.2 Controversy and beyond

After writing the 1974 paper, Grishchuk started working more exclusively with gravitational waves (a notable exception is the paper Grishchuk & Zeldovich (1978), in which the authors show that if we live in a 'typical' part of the Universe, there are no significant density perturbations on scales

bigger than the horizon. This is called the *Grishchuk-Zel'dovich effect*). In 1990, he showed that gravitons created in the early Universe should be in so-called *squeezed states* - quantum-mechanical states where the variance of, e.g., the momentum is very small, making the variance of the position very large due to the uncertainty relation.

It does seem that Grishchuk had no quarrel with any of the 'standard results' of the cosmological community. That is, until 1994, when he published the paper "Density perturbations of quantum-mechanical origin and anisotropy of the microwave background". In it, he followed the generation and evolution of density perturbations through a model universe consisting of an initial stage (which could be inflationary) followed by a radiation stage, and finally a matter stage. He compared the final density perturbations with the final gravitational wave perturbations, and concluded that the standard result claiming that the tensor contribution to the large-scale CMB anisotropy should be negligible compared to the scalar contribution, is wrong - they should be approximately equal, with a slightly higher contribution from the gravitational waves. He followed up with the paper Grishchuk (1995), which was a comment on a paper by Parker and Zhang, whose error, in Grishchuk's eyes, was the use of the tensor-to-scalar ratio predicted by inflation.

This naturally incited reactions. Deruelle & Mukhanov (1995) claimed that Grishchuk's matching conditions between the three stages were incorrect. Caldwell (1996) agreed with Deruelle & Mukhanov, and commented on some other errors he claimed Grishchuk had made. Grishchuk restated his criticisms in the appendix of Grishchuk (1996), but did not say much about what he thought was wrong about the claims of Deruelle & Mukhanov and Caldwell.

Some time later, the paper Martin & Schwarz (1998a) came out, with the aim of clearing up the controversy. The authors said that the previous criticisms of Grishchuk's work were wrong, but also that Grishchuk was wrong, albeit for different reasons than those given by Deruelle & Mukhanov and Caldwell. Another paper (Gotz, 1998) came out shortly after, addressing many of the same issues discussed by Martin & Schwarz, again claiming that Grishchuk was wrong. Grishchuk responded to these claims with Grishchuk (1998), and Martin & Schwarz replied to his response (Martin & Schwarz, 1998b), claiming the controversy to be settled.

Grishchuk never replied to this last paper by Martin & Schwarz. However, he did continue, and still continues, to claim the erroneousness of the standard inflationary result concerning the tensor-to-scalar ratio, though after the controversy, he has emphasised more strongly the importance of choosing the correct initial conditions, compared to, e.g., his 1994 paper. This is evident in, e.g. Grishchuk (2005), which is criticised in Lukash (2006).

Grishchuk has written extensively about the possibility of observing gravitational waves directly, and recently, it seems he has turned to the task of detecting the traces of primordial gravitational waves in the CMB: in Zhao et al. (2009a) (coauthored by Grishchuk), the authors did a likelihood analysis of the *WMAP-5 IE* data and concluded that a model with a tensor-to-scalar quadrupole ratio $R = 0.24$ would be the best fit to the data. They further claimed that the IE data would be more reliable than the BB data in the search for primordial gravitational waves. This article was followed by Zhao et al. (2009b) (also coauthored by Grishchuk), which extended the likelihood analysis using a Monte-Carlo approach, and including more of the cosmological parameters in the analysis.

We will look at the above papers closer in the following chapters. In chapter 8, we will examine the controversy regarding the tensor-to-scalar ratio, and try to pass some judgements on the various claims made. In chapter 9, we examine the results of Zhao et al. (2009b), and try to improve the analysis in several respects.

Chapter 8

The Controversy

This chapter is devoted to the investigation of the controversy that followed the paper "Density perturbations of quantum-mechanical origin and anisotropy of the microwave background" (Grishchuk, 1994). This paper contests the validity of the inflationary predictions, especially the one holding that the tensor contribution to the CMB anisotropy should be negligible compared to the scalar contribution.

We will present a summary of the controversy, focusing on what is perceived as the most important points. In between, and at the end, we will make comments that propose a way of understanding the contesting claims.

8.1 Grishchuk's initial arguments

In this section, we present the arguments put forth by Grishchuk, claiming that the predictions of standard inflationary theory are wrong. The main argument is found in Grishchuk (1994), and will be presented in subsection 8.1.1, but we will also present additional arguments found in Grishchuk (1995) and Grishchuk (1996) in the two following subsections.

8.1.1 The tensor-to-scalar quadrupole ratio

This subsection is devoted to reviewing the 1994 paper as consisely as possible, but with as much detail as is needed for the further analysis. Some results also draw from Grishchuk (1993a), especially those concerning the predictions for gravitational waves (as opposed to scalar perturbations).

Definitions and conventions

We must note, first of all, that Grishchuk uses the *synchronous* gauge. This complicates the comparison of Grishchuk's results to the standard results presented in Part I, but Grishchuk eventually translates his results into observable predictions for the CMB power spectrum, which are easy to compare with the predictions for this observable stated in chapter 6.

We define

$$\gamma(\eta) = 1 - \frac{\dot{\alpha}}{\alpha^2}, \tag{8.1}$$

where α, from eq. (3.11), is the inverse comoving Hubble radius, and γ can also be written as $\gamma(t) = -(H'/H^2)$. $\gamma = 2$ for a radiation-dominated universe, $\gamma = 3/2$ for a matter-dominated universe, and $\gamma = 0$ for a de Sitter universe.

Looking back at eq. (3.1), and recalling that the synchronous gauge is defined by $\phi = B = 0$, we will work with scalar perturbations h, h_l such that $2\psi = hQ$, and $-2E_{,i,j} = h_l k^{-2} Q_{,i,j}$. Thus, in this formalism, the perturbations h and h_l are the *amplitudes* of the perturbations, while the function Q contains their transformation properties. It might seem we have switched the two variables ψ and E for three new ones: h, h_l, and Q. However, the function Q must satisfy

$$Q^{,i}{}_{,i} + k^2 Q = 0, \tag{8.2}$$

thus reducing the degrees of freedom by one.

We will not be concerning ourselves with the perturbations to the energy-momentum tensor, as it relates only tangentially to the main argument of this section. However, we will in the following at least need to know the form of these perturbations:

$$\begin{aligned} T^0{}_0 = -\rho^{(0)} - \frac{1}{a^2}\delta\rho Q, \quad T^0_i = \frac{1}{a^2}\dot{\xi} Q_{,i}, \quad T^i{}_0 = -\frac{1}{a^2}\dot{\xi} Q^{,i} \\ T^k_i = p^{(0)}\delta^k{}_i + \frac{1}{a^2}(\delta p + \delta p_l) Q \delta^k{}_i + \frac{1}{a^2 k^2}\delta p_l Q^{,k}{}_{,i}. \end{aligned} \tag{8.3}$$

Evolution equations must be found for all of the four perturbations above, but, again, this will not be central in the following.

We will assume that the Universe underwent three stages of evolution: an initial stage i, governed by a scalar field, as in chapter 4, a subsequent radiation dominated stage e, and a final matter dominated stage m. This is a pretty decent approximation to what we believe our Universe to be like (see the discussion on the evolution of the Universe in chapter 2), especially in this context where we are just looking at order-of-magnitude estimates.

The scale factor during these three stages can be written as

$$a = l_0|\eta|^{1+\beta}, \tag{8.4}$$

where l_0 is some normalisation constant. In a de Sitter universe, $\beta = -2$. For the e and m stages, $\beta = 0$ and $\beta = 1$, respectively. The times of transition between the three stages are denoted η_1 (i-e) and η_2 (e-m).

Scalar perturbations at the i stage

Using the Einstein equations for the energy-momentum tensor of the scalar field, we find that h_l satisfies

$$\ddot{h}_l + 2\alpha\dot{h}_l - k^2 h = 0, \tag{8.5}$$

and thus is determined once the perturbation h is determined, a result which holds for the energy-momentum perturbations as well. The h perturbation, in turn, satisfies

$$\begin{aligned} &\dddot{h} + \ddot{h}\left(3\alpha\gamma - \tfrac{\dot{\gamma}}{\gamma}\right) \\ &+\dot{h}\left(k^2 - 2\dot{\alpha} + 2\gamma\alpha^2 - \tfrac{\dot{\alpha}}{\alpha}\tfrac{\dot{\gamma}}{\gamma} - \tfrac{d}{d\eta}\left(\tfrac{\dot{\gamma}}{\gamma}\right)\right) \quad + k^2\alpha\gamma h = 0, \end{aligned} \tag{8.6}$$

which is a third-order differential equation. It has a trivial solution $h = 0$, but this is equivalent to the solution

$$h = C\frac{\alpha}{a}. \tag{8.7}$$

The reason for this is that, as we mentioned in chapter 3, the synchronous gauge is not completely determined by the conditions $\phi = B = 0$ - there are infinitely many gauges satisfying this condition, and by performing a coordinate transformation

$$\bar{\eta} = \eta - \frac{C}{2a}Q, \quad \bar{x}^i = x^i - \frac{C}{2}Q^{,i}\int\frac{d\eta}{a}, \tag{8.8}$$

we are still situated in the synchronous gauge, with the redefinitions

$$h \to \bar{h} = h + C\frac{\alpha}{a}, \quad h_l \to \bar{h}_l = h_l + Ck^2\int\frac{d\eta}{a}. \tag{8.9}$$

Thus, since $h = 0$ is a solution to eq. (8.6), (8.7) is also a solution to this equation due to our freedom to choose our synchronous coordinate system among the set defined by eq. (8.9). It is, however, possible to construct so-called 'residual gauge-invariant quantities', meaning quantities defined

within the synchronous gauge, and which do not change under the transformation (8.8):

$$u \equiv \dot{h} + \alpha\gamma h, \quad v \equiv \dot{h}_l - \frac{1}{\alpha}k^2 h. \tag{8.10}$$

In terms of u, with the trivial equation eliminated, eq. (8.6) becomes

$$\ddot{u} + \dot{u}\left(2\alpha\gamma - \frac{\dot{\gamma}}{\gamma}\right) + u\left(k^2 - 2\dot{\alpha} - \alpha\frac{\dot{\gamma}}{\gamma} - \frac{d}{d\eta}\left(\frac{\dot{\gamma}}{\gamma}\right)\right) = 0. \tag{8.11}$$

Finally, in terms of the variable μ, defined through

$$u = \frac{\alpha\sqrt{\gamma}}{a}\mu, \tag{8.12}$$

eq. (8.11) gives

$$\ddot{\mu} + \mu\left(k^2 - \frac{\ddot{(a\sqrt{\gamma})}}{a\sqrt{\gamma}}\right) = 0, \tag{8.13}$$

which looks very much like eq. (7.2). Indeed, for the scale factors (8.4) that we are working with, γ is constant, and eq. (8.13) becomes identical to the analogous expression for gravitational waves, and we will see the same mechanism of superadiabatic amplification as we saw in the gravitiational wave case. For these scale factors, the solution to the above equation is

$$\mu(\eta) = (k\eta)^{1/2}\left(A_1 j_{\beta+\frac{1}{2}}(k\eta) + A_2 j_{-(\beta+\frac{1}{2})}(k\eta)\right), \tag{8.14}$$

where the j's are the spherical Bessel functions, as before. This equation is valid for both μ^{gw} (at any stage) and μ (at the i stage), and so the solution for the perturbation h at the i stage becomes

$$h(\eta) = \frac{\alpha}{a}\int_{\eta_0}^{\eta} d\eta\mu\sqrt{\gamma} + \frac{\alpha}{a}C_i, \tag{8.15}$$

where η_0 is some initial time and the term involving C_i is the trivial solution mentioned above (there will be similar trivial solutions at the e and m stage, hence the subscript).

Since h_l (and all the other perturbations) is determined by h, we have three unknown constants A_1, A_2, and C_i which we need to determine, and they must be determined through initial conditions. However, C_i represents a 'residual' gauge-freedom, and it can be fixed by choosing a specific constant-time hypersurface. We will fix our coordinate system at the m stage to be the coordinate system comoving with the matter fluid, and thus eliminate this residual gauge-freedom.

Scalar perturbations and the e and m stages

We have found an integral solution of the scalar perturbation h at the i stage, and must do the same for the stages dominated by a perfect fluid - the e and m stages.

Not specifying the nature of the perfect fluid yet, we get an equation similar to eq. (8.13):

$$\ddot{\nu} + \nu\left(k^2 c_l^2 - W(\eta)\right) = 0. \tag{8.16}$$

Here, we are working with a new variable ν, defined through

$$u = \frac{\alpha\sqrt{\gamma}}{a} c_s \nu, \tag{8.17}$$

where u is defined in eq. (8.10), and we have also defined c_l and c_s as

$$\frac{\delta p}{\delta \rho} = c_l^2, \quad \frac{\dot{p}^{(0)}}{\dot{\rho}^{(0)}} = c_s^2. \tag{8.18}$$

The function $W(\eta)$ depends on the scale factor and its derivatives, along with c_s, but its exact form is not very interesting and we omit it here.

The radiation-dominated stage

At this stage, even though we have said that our scale factor is of the form (8.4), we must write it more generally as

$$a(\eta) = l_0 a_e(\eta - \eta_e), \tag{8.19}$$

where a_e and η_e are constants, to be determined by the continuous joining of the e stage to the i and m stages. We have not yet said which quantities should be joined continuously; this will be explained later.

We again find that all perturbations can be expressed in terms of h, so we focus only on finding the expression for this perturbation. Its solution is very similar to eq. (8.15):

$$h(\eta) = \frac{\alpha}{a}\int_{\eta_1}^{\eta} d\eta \nu + \frac{\alpha}{a} C_e, \tag{8.20}$$

where the notable differences are the replacements $\mu\sqrt{\gamma} \to \nu$, $\eta_0 \to \eta_1$, and $C_i \to C_e$. The last of these we already anticipated - it's the trivial solution, but now at the radiation-dominated stage. The second replacement reflects that we must choose the initial conditions of the e stage in such a way that

it is continuously joined to the i stage at the transition time η_1. In order to really make sense of eq. (8.20), we need an expression for ν. At the e stage, the potential $W(\eta)$ vanishes, and ν becomes

$$\nu = B_1 e^{-i\frac{k}{\sqrt{3}}(\eta-\eta_e)} + B_2 e^{i\frac{k}{\sqrt{3}}(\eta-\eta_e)}, \tag{8.21}$$

so that we again end up with three constants total to be determined at the e stage. These will be related to the constants at the i stage through the joining of the two stages.

The matter-dominated stage

We write the scale factor here as

$$a(\eta) = l_0 a_m (\eta - \eta_m)^2. \tag{8.22}$$

This time, we cannot completely determine h_l from h; the solutions become

$$h = C_1, \quad h_l(\eta) = \frac{1}{10} C_1 k^2 (\eta - \eta_m)^2 - \frac{1}{3} C_2 \frac{(\eta_2 - \eta_m)^3}{(\eta - \eta_m)^3}. \tag{8.23}$$

With these solutions, we have fixed the coordinate system to be the one comoving with the matter fluid, as we said we would earlier. Otherwise, we would have had terms involving some constant C_m similar to C_i and C_e above. This choice of coordinate system corresponds to the choice $C_m = 0$.

Joining the three stages

We now have solutions for the perturbations at the three stages, and must somehow join these together, relating the constants at the m stage (C_1, C_2) to the quantities at the i stage.

We will join the three stages so that h and h_l are continuous over the transitions. In addition, the scale factor must be continuous always, and v and μ will be continuous over the first transition, although we have not really been working with the quantity μ at the e stage, using the quantity ν instead.

There is a problem with the continuity of h (and h_l): At the i stage, these quantities depend on γ through eq. (8.13). However, $\gamma = (2+\beta)/(1+\beta)$ at the i stage ($= -2$ if the i stage is a de Sitter space), while $\gamma = 2$ at the e stage, so that γ experiences a finite jump at the transition and the time derivative of γ becomes ill-defined.

We deal with this problem by parametrising γ. The exact way of doing this is not important here, but it must be done in such a way that at the time $\eta_1 - \sigma$, where η_1 is the transition time, $\gamma \to (2+\beta)/(1+\beta)$, while at $\eta_1 + \sigma$, $\gamma \to 2$. This must also be done in such a way as to make γ continuous. We then let $\sigma \to 0$. Such a treatment will leave $\dot{\mu}$ discontinuous, but this function never enters anywhere by itself - it always appears with other terms which cancel the discontinuity.

With all the above in place, we let the joining begin. The calculations are long-winded and the resulting expressions for C_1 and C_2 are not enlightening, so we will avoid writing them down before some approximations are made.

First, the expressions for C_1 and C_2 depend on the value of γ at the transition η_1. This value is only well-defined if we are either at the *beginning* of the transition, where $\gamma = (2+\beta)/(1+\beta)$, or at the *end*, where $\gamma = 2$. We will assume we are at the end, and that γ has become a constant, so that $\dot{\gamma} = 0$.

Further, we will limit our discussion to long wavelengths, since these are the ones being superadiabatically amplified.

Lastly, we will set $A_2 = 0$ in eq. (8.14) because this corresponds to the decaying mode in eq. (7.3), which is subdominant compared to the growing mode. This approximation is justified by the fact that averaged over initial phases, the growing solution will always dominate.

Using these approximations, we get

$$\mu(\eta_1) \approx \frac{A_1}{2^{\beta+\frac{1}{2}}\Gamma(\beta+\frac{3}{2})}(k\eta_1)^{\beta+1}, \tag{8.24}$$

where β is its value at the i stage. Further, we get

$$C_1 \approx \frac{1}{2a_1}\sqrt{2}\mu(\eta_1), \quad C_2 \approx -\frac{6}{5a_1}\left(\frac{k}{k_c}\right)^2\sqrt{2}\mu(\eta_1), \tag{8.25}$$

where $k_c \equiv \sqrt{3}/(\eta_2 - \eta_e)$. The long-wavelength approximation means, in this context, $k < k_c$. Let us now compare the value of C_1 $(= h)$ to the amplitude of gravitational waves upon entering the m stage (Grishchuk, 1993a):

$$h^{gw}(\eta_2) \approx \frac{3\sqrt{3\pi}}{\sqrt{2}}\frac{1}{a_1}\mu^{gw}(\eta_1). \tag{8.26}$$

If we assume that the absolute value of μ^{gw} is not too different from that of μ, we see that $h^{gw}/h \approx 3\sqrt{3\pi}$, and there is no radical difference between the amplitudes of density perturbations and gravitational waves, independent of the parameters of the i stage. Defining a 'characteristic' amplitude $h_c(k) \sim kl_{Pl}C_1$, where $l_{Pl} = \sqrt{G}$ is the *Planck length*, we find that the characteristic amplitude $h_c(k) \sim (l_{Pl}/l_0)k^{\beta+2}$.

Initial conditions

The initial conditions we use are essentially the same as those derived in chapter 4, with μ replacing φ in the discussion of section 4.1.3, since, as long as γ is constant, these are the same quantities, as is evident from comparing eq. (8.13) and eq. (4.15) (the treatment in Grishchuk's article is more complete, as it considers the 'complete' field - both the scalar field and the scalar metric perturbations, but the results are essentially the same, since, as we said, the metric perturbations are negligible during inflation). Using the initial conditions derived in this way, one can find the value

$$A_1 = -\frac{i}{\cos\beta\pi}\sqrt{\frac{\pi}{2}}e^{i(k\eta_0+\frac{\pi\beta}{2})} \tag{8.27}$$

for the constant in eq. (8.24). The most important thing to note about this equation is that it does not contain any 'smallness parameters' such as γ, so that the earlier assumption about the absolute value of μ not being severely different from μ^{gw} are valid.

The quadrupole tensor/scalar ratio

We now want to relate our solution for the scalar perturbations to the intensity (temperature) anisotropies today. The approach taken in Grishchuk (1994) is a bit different from that taken in chapter 6, so we will only give the results that we need for the further analysis.

First, only the derivatives of the scalar perturbations contribute to the temperature anisotropy in the synchronous gauge. Therefore, since h at the onset of the matter dominated epoch is a constant, it will not contribute to the anisotropy, which means only h_l will contribute.

Second, the C_l^{II}'s are proportional to $|C_1|^2$.

Using the above results and statements, one may derive the formula

$$\frac{C_l^{T,II}}{C_l^{II}} \approx \frac{(l+1)(l+2)(2l+1)^2}{2l(l-1)}. \tag{8.28}$$

One then gets for the quadrupoles

$$\frac{C_2^{T,II}}{C_2^{II}} \approx 75. \tag{8.29}$$

This is in disagreement with the results obtained in chapter 6, for example fig. 6.1, where, certainly, the tensor-to-scalar quadrupole ratio is considerably less than unity.

This concludes our review of Grishchuk (1994), which contained his main argument. He had, however, more to say on the subject of the theory of inflation, and we will briefly recount what he said below.

8.1.2 The constancy of ζ

We saw in chapter 4 that an important part of the predictions for the power spectra at the end of inflation was the use of the long-wavelength 'conservation law' for the quantity ζ, defined in eq. (4.12). But what does this conservation law really tell us?

In terms of μ, defined in (8.12), the potential Ψ becomes

$$\Psi = \frac{1}{2k^2}\alpha\gamma\frac{d}{d\eta}\left(\frac{\mu}{a\sqrt{\gamma}}\right). \tag{8.30}$$

Using eq. (4.11) with the above version of Ψ, one gets

$$\frac{1}{a^2\gamma}\frac{d}{d\eta}\left(a^2\gamma\frac{d}{d\eta}\left(\frac{\mu}{a\sqrt{\gamma}}\right)\right) + k^2\frac{\mu}{a\sqrt{\gamma}} = X, \tag{8.31}$$

where X is an integration constant. Using the definition of ζ to rewrite it in terms of μ, one gets

$$\zeta = \frac{1}{2k^2}\frac{1}{a^2\gamma}\frac{d}{d\eta}\left(a^2\gamma\frac{d}{d\eta}\left(\frac{\mu}{a\sqrt{\gamma}}\right)\right) = \frac{X}{2k^2} - \frac{\mu}{2a\sqrt{\gamma}}, \tag{8.32}$$

where the last equality follows from eq. (8.31). Then, we take the long-wavelength limit of this equation, giving $\zeta \approx X/2k^2$, and use the additional fact that the left hand side of eq. (8.31) is the same as the left hand side of eq. (8.13), only slightly rewritten. This means that the integration constant X must in fact be zero, resulting in $\zeta = 0$. Thus, the 'conservation' of ζ is just an empty statement, saying that $0 = 0$.

8.1.3 The large amplification of Ψ

One of the features of inflationary theory is that it predicts that the superhorizon modes of the potential Ψ will be amplified during its crossing from the inflationary stage to the matter-dominated stage: Consider eq. (4.12). By assuming that superhorizon modes of Ψ are constant at the initial and matter-dominated stages (which is a pretty accurate assumption), ζ reduces to

$$\zeta \approx \frac{5+3w}{3(1+w)}\Psi, \tag{8.33}$$

which, when the constancy of ζ for superhorizon modes is used, translates into

$$\Psi(\eta_m) \approx \frac{1}{1+w(\eta_i)}\Psi(\eta_i), \tag{8.34}$$

and since w is typically close to -1 during the initial stage, this means that the potential has been amplified during the transition from the initial stage to the matter-dominated stage (all the time, of course, assuming that the wavelength is so large that the mode does not enter the horizon during the radiation dominated stage. Anyway, a similar amplification will occur for such modes).

However, by rewriting eq. (4.12) as

$$\zeta = -\frac{H^2}{aH'}\left(\frac{a}{H}\Psi\right)' \tag{8.35}$$

and integrating, remembering that $\zeta = \zeta_0$ is a constant, we find

$$\Psi = \zeta_0\left(1 - \frac{H}{a}\int_{t_i}^{t} dta\right) + C\frac{H}{a}. \tag{8.36}$$

The time t_i is the time of horizon exit. Similarly, we denote the time of horizon entry by t_f. We stick with the case of an initial inflationary stage followed by a radiation-dominated, or eventually matter-dominated, stage, with the scale factors[1]

$$a_i = a_1 t^{p_1}, \quad a_f = a_2(t-t_*)^{p_2}. \tag{8.37}$$

The quantities a_2 and t_* will be expressible in terms of the other quantities by demanding continuity of the scale factor. The power p_1 will typically be much larger than one, while $p_2 = 1/2$ or $2/3$, depending on what final epoch we are considering. We denote the time of transition by t_1. The initial and final values of Ψ become

$$\Psi(t_i) = \zeta_0 + C\frac{H(t_i)}{a(t_i)}, \quad \Psi(t_f) = \zeta_0 I + C\frac{H(t_f)}{a(t_f)}, \tag{8.38}$$

where

$$I \equiv 1 - \frac{H(t_f)}{a(t_f)}\int_{t_i}^{t_f} dta \approx \frac{1}{p_2+1}, \tag{8.39}$$

the last approximation being valid since $a(t_f) \gg a(t_i)$, $p_1 \gg p_2$ and $a(t_1)H(t_1) \gg k$ (since t_1 presumably is later than t_i). This leads to

$$\Psi(t_f) = \Psi(t_i)I + C\left(\frac{H(t_f)}{a(t_f)} - \frac{H(t_i)}{a(t_i)}I\right) \approx \frac{1}{p_2+1}\Psi(t_i) - C\frac{H(t_i)}{a(t_i)}\frac{1}{p_2+1}. \tag{8.40}$$

[1] This form of the scale factor, used in Grishchuk (1996), differs from the form (8.4), used in Grishchuk (1994), because the 1996 paper is in part a reply to Deruelle & Mukhanov (1995), in which this form of the scale factor is used.

Here, one may set $C = 0$, but in any case, the term involving C is smaller than or of the same order as the first term. Setting $C = 0$ gives

$$\Psi(t_f) \approx \frac{1}{p_2 + 1}\Psi(t_i), \tag{8.41}$$

showing that there is no large amplification of Ψ in the transition from the inflationary stage to the radiation-or-matter-dominated stages.

8.2 Reactions

The last section contained Grishchuk's main criticisms concerning the inflationary predictions. As could be reasonably expected when a challenge is issued to well-established ideas, there were reactions to these criticisms, as described in chapter 7. As also mentioned there, Martin & Schwarz (1998a) showed that the treatments of Deruelle & Mukhanov (1995) and Caldwell (1996) were incorrect, and provided their own criticism of Grishchuk's work. As Martin & Schwarz address all of the points of the previous section, we will assume that they are right in saying that Deruelle & Mukhanov and Caldwell are wrong, and follow their line of reasoning.

If inflationary theory is correct, then each of the above subsections must have some flaw in their reasoning, and this is what is explored below.

8.2.1 The constancy of ζ

Even though the integration constant X in eq. (8.31) is zero, it does not follow that $\zeta = 0$ in the long-wavelength limit. Indeed, if we use the correct value $X = 0$ in eq. (8.32), we see that $\zeta = -\mu/2a\sqrt{\gamma}$, which is nonzero, and expanding the right hand side to leading order, one finds that it is indeed a constant. In this expansion, however, one also gets a decaying mode in addition to the constant term. In order to make use of the constancy of ζ, one should always keep in mind that one neglects this decaying mode.

8.2.2 The large amplification of Ψ

We here consider the scale factor (8.37). Let us write out eq. (8.36) at the inflationary stage:

$$\Psi(t < t_1) = \frac{\zeta_0}{1 + p_1}\left(1 - \left(\frac{t_i}{t}\right)^{p_1+1}\right) + \frac{p_1}{a_1 t^{p_1+1}}C. \tag{8.42}$$

In the previous subsection, it was noted that in order to use the constancy of ζ_0, one must neglect the decaying mode, which here means that the terms proportional to $1/t^{p_1+1}$ must vanish. This fixes the constant C and gives the pre-transition value

$$\Psi(t < t_1) = \frac{\zeta_0}{1 + p_1}, \tag{8.43}$$

which gives $\Psi = 0$ in the limit $p \to \infty$. This is correct, as this limit corresponds to the de Sitter space, and we stated in chapter 4 that there were no scalar perturbations in de Sitter space. If, alternatively, one uses Grishchuk's prescription $C = 0$, one gets from eq. (8.38) that Ψ is a constant independent of p_1 during inflation. Using the corrected value for Ψ, one again ends up with the result that the superhorizon potential is greatly amplified during the transition from the inflationary to the matter (or radiation) stage.

8.2.3 The tensor-to-scalar ratio

We remember that, in Grishchuk's treatment, the quantity μ was said to be continuous over the transition from the i stage to the e stage. However, we know from the discussion on ζ that in the long-wavelength limit, which Grishchuk considered in his 1994 paper, $\zeta = -\mu/2a\sqrt{\gamma}$, which is a constant to leading order. This means that

$$\frac{\mu}{a\sqrt{\gamma}}(\eta_1 - \sigma) \simeq \frac{\mu}{a\sqrt{\gamma}}(\eta_1 + \sigma), \tag{8.44}$$

and since $\gamma(\eta_1 - \sigma) \neq \gamma(\eta_1 + \sigma)$, we cannot have that $\mu(\eta_1 - \sigma) = \mu(\eta_1 + \sigma)$. Specifically, we must have

$$\mu(\eta_1 + \sigma) \simeq \frac{\mu(\eta_1 - \sigma)}{a(\eta_1 - \sigma)\sqrt{\gamma(\eta_1 - \sigma)}} a(\eta_1 + \sigma)\sqrt{\gamma(\eta_1 + \sigma)} = \mu(\eta_1 - \sigma)\sqrt{\frac{2}{\gamma_i}}, \tag{8.45}$$

where γ_i is the value of γ at the i stage (typically close to zero), and $\gamma_e = 2$. We see the appearence of an 'amplifying' factor $1/\sqrt{\gamma}$, which is responsible for the largeness of scalar perturbations upon entry to the radiation-dominated stage. Claiming that μ is constant across the transition from the i to the e stage, as Grishchuk did, leads to the erroneous result that the scalar and tensor amplitudes should be roughly equal.

8.3 Backreactions and a temporary end to the controversy

Grishchuk responded to Martin & Schwarz' criticism with Grishchuk (1998), and we shall look at his response and, finally, their comment to his response (Martin & Schwarz, 1998b).

8.3.1 Grishchuk's response

Let us define a new parameter $\bar{\mu}$ through

$$\bar{\mu} = \frac{\mu}{\sqrt{\gamma}}. \tag{8.46}$$

Grishchuk acquiesces that it is indeed this variable, and not μ, that is continuous over the *i-e* transition. He says that this was actually a typo, and that the continuity of μ was never used in practice when joining the three stages together, so they have no bearing on the final results.

He further qualifies some of his earlier statements: The conservation law for ζ is empty, he now says, in the sense that it does not add anything new that we did not already know by simpler means, since it can be written as $\zeta \approx -\bar{\mu}/2a$, and the function $\bar{\mu}$ has been analysed in Grishchuk (1994).

Then, concerning the 'big amplification' of Ψ, Grishchuk this time around rederives and agrees to the standard result that $\Psi(\eta_m)/\Psi(\eta_i) \sim 1/(1+w_i)$ where w_i is the equation of state at the i stage (i.e. close to -1). However, he says it is a misinterpretation to say that Ψ is amplified - rather, it is only at the m stage that Ψ reaches the amplitude that, say, the gravitational waves have had since the beginning. This must, in the end, be determined by the initial conditions, he says. He restates the result from Grishchuk (1994) that the 'characteristic amplitude' of the scalar metric perturbation is $h_c(k) \approx (l_{Pl}/l_0)k^{2+\beta}$.

8.3.2 Closure by Martin & Schwarz

The short paper Martin & Schwarz (1998b) contains replies to Grishchuk's comment. The introduction of $\bar{\mu}$, they say, does not add anything to the problem. In particular, they disbelieve Grishchuk's claim that the continuity of μ was a misprint, and claim that it was indeed this assumption that led to the error in the 1994 paper.

They end their paper by saying, "We consider this controversy to be settled now".

8.4 Interlude

In this section, we will try to comment on the discussion so far. First, we may note that one of Grishchuk's most repeated complaints about inflation is its "absurd" prediction that the density perturbation will be "infinitely" big at the beginning of its evolution. We will comment on this later.

Ultimately, one's feelings of absurdness does not matter much in science. Rather, we should ask: Is Grishchuk correct? To answer this question, it may be worthwhile to sum up the above discussion to get an overview of the situation.

In order to do this, we must also try to clear up some of the points of conflict:

When it comes to the argument over the constancy and nonemptiness of ζ, we must assert that Grishchuk initially either a) believed that in the long-wavelength limit, the quantity ζ was actually zero, and was proved wrong and admitted it, or b) that he articulated his case very badly in the 1994 paper, in which he wrote that the constancy of ζ reduced to the empty statement $0 = 0$. In the 1998 paper, on the contrary, he says that the constancy of ζ is empty in the sense that it gives nothing more than what one could arrive at using simpler methods.

Concerning the large amplification of Ψ, only the option that Grishchuk initially believed something which turned out to be erroneous seems likely. He explicitly says, in Grishchuk (1996), that "There is nothing like tremendous jumps of $[\Psi]$ at the transition point", after deriving eq. (8.41), which, in mathematical form, says that Ψ is not amplified. However, as we saw, in his 1998 paper, he accepts eq. (8.34), but will not accept that this is a large amplification. Rather, he says it means that Ψ was very close to zero initially, and has just recently acquired a moderate value. This seems to be just a game with words - regardless of the initial value of Ψ, it *has* been amplified, but of course its absolute value compared to, say, the tensor perturbation, must be determined by other means, with which Martin & Schwarz also agrees.

Finally, in the argument leading up to eq. (8.29), it seems there is an unresolved issue: Did Grishchuk make use of the continuity of μ, as Martin & Schwarz claims, or did he not, as he himself claims? This seems easily verifiable enough: Join the solutions at the i, e and m stages, and see if

we get the same result as Grishchuk did for C_1 and C_2 without using the continuity of μ. When doing this, we find that this is indeed possible, and that the continuity of μ was *not* used in joining the three stages.

However, the story does not end there: As we noted at page 63, in order to actually be able to make sense of the expressions for C_1 and C_2, we must determine whether we are at the beginning or end of the $i-e$ transition. As we also stated, we decided to be at the end, where $\gamma = 2$. This means that eq. (8.14) is no longer applicable, unless μ is continuous over the transition (which both Martin & Schwarz and Grishchuk agree that it isn't). However, eq. (8.24) is the long-wavelength approximation of eq. (8.14), where the decaying mode has been neglected. In treating this equation as valid at the end of transition, one must assume the continuity of μ. This means that Grishchuk actually does assume this continuity, and that, in light of eq. (8.45), the correct form of eq. (8.14) is

$$\mu(\eta_1) \approx \frac{A_1}{2^{\beta-\frac{1}{2}}\Gamma(\beta+\frac{3}{2})\sqrt{\gamma_i}}(k\eta_1)^{\beta+1}, \tag{8.47}$$

which, since γ_i is close to zero, should give some radically different results. Most notably, we find, since $C^{II} \propto |C_1|^2$, and C_1 is given by eq. (8.25), that instead of eq. (8.29), the tensor-to-scalar quadrupole ratio is $\approx 75\gamma_i$, which is more like the inflationary predictions, in that it actually depends on the 'smallness' parameter γ_i (see eq. (4.33), which admittedly is not the quadrupole ratio, but since it depends on ϵ, and the quadrupoles more than other multipoles reflect the primordial initial conditions, we expect that the quadrupole ratio also should depend on ϵ, or some other smallness parameter, like γ_i. The exact numerical factor depends on where one fixes γ_i, and we have not derived any such factor for the quadrupole ratio in this text).

8.4.1 New initial conditions?

We saw that in the 1994 paper, Grishchuk argued that the 'characteristic amplitude' of the scalar perturbations, proportional to C_1, was given by $h_c \sim k^{\beta+2}$. We also saw that after having adopted the correct transition conditions, Grishchuk restated this result, and claimed it to be valid. But since $h_c \propto C_1$, and C_1 is given by eq. (8.25), and μ has changed from eq. (8.24) to (8.47), then in order for h_c to stay unchanged (or at least independent of γ_i), we need the constant A_1 to change as well: $A_1 \to A_1\sqrt{\gamma_i}$. This can only come about if the initial conditions are different, so we should expect Grishchuk to adopt a different initial condition scheme than that used in the 1994 paper (which is basically the same as the one used in Part I). We

will examine this below, but it is worth keeping in mind that *if* Grishchuk adopts a different initial condition scheme than he previously used, then there is a lack of consistency in his treatment which stays unexplained. However, we should evaluate the claims by themselves, not their motivations, and that is what we will do.

8.4.2 Tying up loose ends

A lot of what we have done in this chapter seems quite different from the treatment of Part I in general, and in chapter 4 in particular. However, there are connections between the two, and we may try to make them in order to understand the discussion about initial conditions better.

First, the connection between the scalar field and the γ parameter, during inflation, is (Grishchuk, 1994)

$$\left(\frac{\phi^{(0)\prime}}{H}\right) = \sqrt{2} M_{Pl} \sqrt{\gamma}. \tag{8.48}$$

Thus, both $\phi^{(0)\prime}$ and γ represent a 'small' quantity during inflation.

Further, the relation between the scalar field perturbation and the parameter μ is

$$\varphi = \frac{M_{Pl}}{\sqrt{2}} \left(\mu - a\sqrt{\gamma} h\right). \tag{8.49}$$

This shows that during the inflationary period, when the parameter γ is supposed to be small, and the metric perturbations are negligible (as we said in chapter 4 that they were), $\mu \sim \varphi$, and in the treatment of Part I, μ is the variable to be treated as a harmonic oscillator as we did φ. This is further highlighted by looking at eq. (8.13), which, when γ is constant, turns into the exact same evolution equation as for φ.

8.5 The proper quantisation of the early perturbations

In this section, we will examine the quantisation scheme adopted by Grishchuk after the controversy with Martin & Schwarz. We will concentrate on Grishchuk (2005), in which his thoughts are spelled out clearly. We will also look at Lukash (2006), which criticises Grishchuk's treatment of the perturbation quantisation. Throughout, we will present the case as pertains to the scheme we outlined in chapter 4, even though it is presented both a bit differently and in more detail in the papers we review.

8.5.1 Grishchuk's treatment

Grishchuk here discusses quantisation in terms of ζ instead of the scalar field, which is just a matter of switching between definitions using eq. (4.25)[2].

Grishchuk starts with the gravitational waves, writing down an expression for the operator $\hat{q}$, which is just the normalised, quantised version of the tensor perturbations $h_{+,\times}$ that we encountered in chapter 4 (we drop the subscripts from now on). 'Normalised' here means that he introduces

$$\bar{h} \equiv \frac{a_0 M_{Pl}}{k} h, \tag{8.50}$$

where a_0 is the scale factor at some initial time η_0 where initial conditions are to be set, so that $\bar{h}$ is the quantity he quantises according to

$$\hat{q}_T \equiv \hat{\bar{h}} = \sqrt{\frac{1}{2}} \frac{a_0}{a} \left(\hat{c} e^{-ik(\eta-\eta_0)} + \hat{c}^\dagger e^{ik(\eta-\eta_0)} \right). \tag{8.51}$$

Let us examine this a bit before moving on. We stated in chapter 4 that the quantisation of the tensor perturbations would follow the exact same procedure as that for φ, so we should compare the above equation to eq. (4.16). Doing so, we note that it is basically the same, except for some constants and the factor $1/a$ in front. The constants are of little concern to us - they are chosen so that the operator $\hat{q}$ satisfies the correct initial conditions, namely those of a harmonic oscillator:

$$\langle 0 | \hat{q}_T^2 | 0 \rangle = \frac{1}{2}. \tag{8.52}$$

(We always have $\hbar = 1$). As mentioned, the above initial condition is set when $\eta = \eta_0$, making $a = a_0$ and everything works out. We should also comment on the $1/a$ factor, which is missing from (4.16). This is easily explained by noting that the tensor perturbation h satisfies the same equation as $\delta\phi$, not φ, and since it is φ that is quantised in chapter 4, and $\varphi = a\delta\phi$, the a enters naturally.

Grishchuk then goes on to density perturbations. Again introducing the normalised variable

$$\bar{\zeta} = \frac{a_0 M_{Pl}}{k} \zeta, \tag{8.53}$$

he writes the operator version of this variable as

$$\hat{q}_S \equiv \hat{\bar{\zeta}} = \sqrt{\frac{1}{2}} \frac{\tilde{a}_0}{\tilde{a}} \frac{1}{\sqrt{\gamma}} \left(\hat{b} e^{-ik(\eta-\eta_0)} + \hat{b}^\dagger e^{ik(\eta-\eta_0)} \right), \tag{8.54}$$

[2]Grishchuk does, however, operate with a slightly different version of ζ, which is related to ours by a factor -1/2.

where $\tilde{a} \equiv a\sqrt{\gamma}$, and $\tilde{a}_0$ is the corresponding initial value for this variable. The usage of $\tilde{a}$ instead of a for density perturbations are justified in Grishchuk (1994): All dynamical equations for scalar-field driven density perturbations can easily be obtained from the gravitational equations by substituting a with $a\sqrt{\gamma}$ and μ_{gw} with μ (compare, for instance, eqs. (8.13) and (7.2)). The physical justification is that whereas the tensor perturbations do not couple to the background scalar field, so that eq. (3.6) is exact, the density perturbations, represented by $\delta\phi$ in our treatment and by μ in Grishchuk's, do couple to the background scalar field, but we have chosen to neglect this coupling in writing eq. (4.10), so that the equations for the density and gravitational perturbations become the same.

Anyway, this justifies the factors of $\tilde{a}$ in the above equation. The extra $1/\sqrt{\gamma}$ factor in front is justified by remembering that eq. (8.51) is the operator expansion of $h = \mu_{gw}/a$, while the above equation is the equation for $\zeta = \mu/a\sqrt{\zeta}$.

However, calculating the expectation value of this (normalised) variable gives $1/(2\gamma)$, not $1/2$ as it should be. This, says Grishchuk, implies that the state annihilated by the operator $\hat{b}$ is in fact *not* the vacuum state of this operator, but rather a multiparticle state. This can be remedied by a so-called *Bogliubov transformation*

$$\hat{b} = u\hat{d} + v\hat{d}^\dagger, \quad \hat{b}^\dagger = u^*\hat{d}^\dagger + v^*\hat{d}. \tag{8.55}$$

By choosing the constants u and v in the correct manner, $\hat{q}_S$ becomes

$$\hat{q}_S = \frac{1}{2}\frac{\tilde{a}_0}{\tilde{a}}\left(\hat{d}e^{-ik(\eta-\eta_0)} + \hat{d}^\dagger e^{ik(\eta-\eta_0)}\right). \tag{8.56}$$

Now, defining the ground state as the state annihilated by the $\hat{d}$ operator instead, the proper initial conditions are recovered. Remembering, then, that $\hat{q}_S$ represented the normalised variable $\bar{\zeta}$, and that $\bar{\zeta} \propto a_0\zeta$, one finds that the spectrum of ζ at horizon crossing, assuming that it does not change much between its initial value and its value at horizon crossing, becomes comparable to, not hugely amplified compared to, the spectrum of the gravitational waves.

8.5.2 Lukash's criticism

In Lukash (2006), the above treatment is criticised, and though the paper itself is several pages long, we translate the point it makes into our context.

The error Grishchuk makes, according to Lukash, is that the variable $\bar{\zeta}$ with which he works is not a properly normalised variable, and so it is an

error to demand that the initial expectation value of $\bar{\zeta}$ should be that of an harmonic oscillator. If we take a look at eq. (8.53) and compare it to eq. (8.50), we find that the exact same normalisation procedure has been employed, even though we noted above that when moving from the treatment of gravitational waves to density perturbations, we should make the replacements $a_0 \to a_0\sqrt{\gamma_0}$. Thus, the correct normalisation of $\bar{\zeta}$ should be eq. (8.53) multiplied by a factor $\sqrt{\gamma_0}$, and we cannot expect the variable Grishchuk has worked with to yield the correct expectation value. Thus, the operators in eq. (8.54) are indeed those that annihilate the vacuum, not a multi-particle state.

Working with the correctly normalised variable, we get the correct result, and by using that the (correctly) normalised variable is related to the original curvature perturbation by $\bar{\zeta} \sim a_0\gamma_0\zeta$, we also recover the standard result, namely that $\langle\zeta^2\rangle$ is proportional to $\gamma_0\langle h^2\rangle$ at horizon crossing - the 'great amplification' (assuming γ_0 is small, which we assume it to be at all times during inflation).

Another comment Lukash makes is that in eq. (8.54), one of the $1/\sqrt{\gamma}$ factors should be replaced with $1/\sqrt{\gamma_0}$, since ζ is proportional to $1/(a\sqrt{\gamma})$, not $1/(a\gamma)$ as one might think, looking at eq. (8.54).

Lukash makes more comments on why Grishchuk's treatment is wrong, but these are formulated in terms of the Lagrangian Grishchuk uses, and to present them, we would have to present more of both Grishchuk's and Lukash's papers. We opt not to do this here, since what we have highlighted so far is sufficient for our purposes.

8.5.3 Comments on the quantisation procedures

We have now presented the two views on this case, and we will try to comment on the correctness of the claims made.

It seems clear that the variable $\bar{\zeta}$ introduced by Grishchuk has been introduced in an inconsistent manner: Why normalise ζ in exactly the same way as h, when in all other cases of going from gravitational waves to density perturbations, we make the substitution $a \to a\sqrt{\gamma}$? This question seems to be unanswered by Grishchuk in his later publications as well - he more or less simply repeats what is said in the paper we analyzed above.

An error in Lukash?

We should also comment on what may be a small error in Lukash's treatment. He said, as we saw, that eq. (8.54) should be corrected by replacing one of the $1/\sqrt{\gamma}$ factors by $1/\sqrt{\gamma_0}$, since otherwise, ζ has the wrong dependence on γ and a. However, we justified both $1/\sqrt{\gamma}$ factors above. Why would we replace one of them by a constant?

In order to comment on this, we may first note that Grishchuk's paper is actually the fourth arXiv version of that paper. In the previous versions, $\tilde{a}$ is replaced by a in eq. (8.54). This gives all the same results as before, with the additional benefit that ζ obtains the correct dependence on γ. It is hard to tell why Grishchuk introduced the $\tilde{a}$ factor instead - perhaps because of a referee comment, or perhaps because of Lukash's paper? Whatever the reason, it enables us to propose the following development of Grishchuk's paper:

Grishchuk first worked with eq. (8.54), but with $\tilde{a}$ replaced by a - which still gave the 'absurd' results that we saw and saw explained above. This form of the $\hat{\bar{\zeta}}$ operator actually seems to be more correct, as we then have a_0 in the numerator and $a\sqrt{\gamma}$ in the denominator, making the expression in parenthesis equal to the operator version of μ (or φ).

Because of some impulse, he then replaced a with $\tilde{a}$. This introduces another factor of $1/\sqrt{\gamma}$. Above (following Grishchuk's argument), we justified one of these factors by appealing to the substitution $a \to a\sqrt{\gamma}$, and the other, we justified by saying that the gravitational wave variable and the curvature perturbation is related through $\zeta = h/\sqrt{\gamma}$. This justification is offered in all versions of Grishchuk's paper, so that regardless of whether one has accounted for the substitution $a \to a\sqrt{\gamma}$ or not, this argument is supposed to hold. However, what is important to realise is that the statement $\zeta = h/\sqrt{\gamma}$ is really just another way of saying that we should make the substitution $a \to a\sqrt{\gamma}$! Thus, by *both* replacing a by $a\sqrt{\gamma}$ *and* introducing the extra factor $1/\sqrt{\gamma}$, we are actually doing the same substitution twice, which leads to erroneous results.

To sum up the above point: eq. (8.54) with a instead of $\tilde{a}$ actually seems like the correct equation: We have the constant a_0 in front, which is due to the normalisation $\bar{\zeta} \sim a_0\zeta$, and a factor $1/(a\sqrt{\zeta})$, which gives the correct dependence of ζ on γ. By introducing $\tilde{a}$ but leaving everything else the same, Grishchuk is being inconsistent. Further, Lukash may be wrong in saying that the solution is to replace one of the γ's by γ_0 - this gives the correct result, but the justification for this replacement is not offered by

Lukash, and it seems that the alternative solution (reverting to a instead of $\tilde{a}$) described above works better.

Summary of Grishchuk's quantisation scheme

What we have seen in this section suggests that eq. (8.54) probably is correct (or at least the version where $\tilde{a}$ is replaced by a), but that Grishchuk's interpretation of it is wrong: this is not the variable that should contain 'half a quantum' in its ground state. Grishchuk blames the operators (or the states), claiming it is not the vacuum they annihilate, but rather a multiparticle state. He *should* have blamed the normalisation of ζ, which is inconsistent with his normalisation of the tensor perturbation h together with the replacement $a \to a\sqrt{\gamma}$. Following the proper normalisation, one again rederives the standard result.

8.6 Chapter summary

All of the above seems to point to the conclusion that Grishchuk has been wrong in his claims. We first noted that there are indeed a large amplification of the scalar perturbations from the scalar field dominated stage to the radiation dominated stage. We then found that the scalar perturbations are also amplified compared to the gravitational waves, which exhibit no such amplification, and that they are not comparable to each other, as Grishchuk claims.

There are some noteworthy points to make at the end of this discussion: Grishchuk does not portray things the way we have done in the above paragraph. According to him, the inflationary community has now realised its error in claiming that the density perturbations are amplified during reheating, and has now moved to the question of initial conditions. Further, he also points to the 'absurd' inflationary claim that the density perturbations were infinitely large from the very beginning.

To respond to this, we must note that there are actually two separate things we might mean when we say 'density perturbations'. We might mean the curvature perturbation ζ, as in chapter 4 or the above discussion about initial conditions, or we might mean the scalar perturbation Ψ (or the equivalent in synchronous gauge), which is what we were mostly talking about in the above discussion about Grishchuk's 1994 paper. It is true that ζ does not acquire its large value during reheating, and that the factor $1/\gamma$ is present from the very beginning, which may be verified by taking the expectation

value of eq. (8.54) divided by a_0. More on this below. But it is also true that Ψ acquires its large value during reheating, as explained in Dodelson (2003) (where it is explained how the large value of ζ gets transferred to Ψ after inflation), and as we saw in the above treatment.

Concerning the evidently large value of ζ from the very beginning, we must say, with Lukash (2006), that, first of all, it is not infinitely big, as Grishchuk repeatedly says. Admittedly, the inflationary period is often referred to as a 'de Sitter' stage, which would make $\gamma = 0$ and the perturbations infinitely large. However, the de Sitter space must be an approximation, since a de Sitter space would correspond to a non-rolling scalar field, which is unviable (Dodelson, 2003). Thus, it is clear that ζ is not infinitely big initially. Lukash shows, in the paper we have just reviewed, that typically, the tensor-to-scalar ratio is of the order 1/5, making the tensor spectrum only five times smaller than the scalar power spectrum. This means that the initial scalar spectrum is neither infinitely nor absurdly big.

Moving back to the question of gravitational waves, we know that ultimately, the only final arbiter in science (ideally) is observation, and the claims made about the relative contribution of tensors in the primordial power spectra should, in principle, be subject to observational tests. In the next chapter, we shall look at some papers Grishchuk coauthored, whose goals (one of their goals, at least), is to test whose predictions are correct.

Chapter 9

Data analysis: Indications of tensor contributions to the CMB

In this chapter, we will look at some of the claims made in Zhao et al. (2009a) and Zhao et al. (2009b), both coauthored by Grishchuk, and we will also try to improve the treatment of the last article. All computations in this chapter have been carried out on the NOTUR Titan cluster.

We will in this chapter ultimately ask the question: What is the cosmological model that best fits the available data? Put another way: We have sets of data and a set of cosmological parameters - given the former, what combination of the latter should we choose? In order to answer this, we begin with a review of Bayesian statistics, and connect it to the context of cosmology. We do this because the articles we will review are based on some knowledge about cosmological applications of Bayes' theorem. Doing this, we will also mention the program CosmoMC and some of its features. However, the discussion concerning CosmoMC is included for later purposes; the authors of Zhao et al. (2009a) and Zhao et al. (2009b) do not use CosmoMC in any way.

9.1 Bayesian statistics and CosmoMC

In this section, we will present the basic theory behind Markov Chain Monte Carlo (MCMC) samplers, and in particular present the sampler which is currently most used for cosmological parameter testing: CosmoMC.

9.1.1 Bayesian inference

This subsection is mostly based on Bernardo (2003) and Christensen et al. (2001).

The world of statistics is divided into two ways of looking at probability and statistics. One is the standard, "toss these dice enough times to see what the probability for two sixes are", frequentist paradigm, while the other is called the *Bayesian* paradigm, which talks about probabilities in the sense "what is the probability that it will rain today?". Central to this paradigm is *Bayes' theorem*:

Let $D = \{x_1, x_2, \cdots x_n\}$ be a general set of data, and let ω be a parametric representation (that is, a vector of several parameters) of the underlying mechanisms that generated this data set. Let, further, K be our knowledge of ω before D was observed. Then, $p(D|\omega)$, here also called the *likelihood* $\mathcal{L}$, is the probability distribution function (PDF) of the data set given a set of parameters ω, while $p(\omega|K)$ - the *prior* - is the PDF of ω given our knowledge of the parameters before observing D. Let Ω be the complete space of parameters, and let A be some assumptions regarding our parametrisation. Then, the *posterior* PDF $p(\omega|D, A, K)$ - that is, the PDF of a given set of parameters given the data, our prior knowledge, and our theoretical assumptions - is given by

$$p(\omega|D, A, K) = \frac{p(D|\omega)p(\omega|K)}{\int_\Omega p(D|\omega)p(\omega|K)d\omega}. \tag{9.1}$$

Usually, we drop the A and K in writing this relation, but it is important to keep in mind that a quantity like $p(\omega)$ is really based on our prior knowledge of, and assumptions about, ω, not on some abstract absolute probability.

We see that the denominator is independent of ω; thus, we may treat this as a normalisation constant, as long as we're not attempting to find the absolute value of $p(\omega_1|D)$ for some parameter set ω_1, but rather this value relative to some other set of parameters ω_2.

9.1.2 Monte Carlo methods, Markov Chains, and the Metropolis-Hastings algorithm

A *Monte Carlo* method is a method which, among other things, aims to answer the question of how to draw a sample from a general, possibly complicated, probability distribution $P(x)$, which might be computationally expensive to evaluate at all points x (McKay, 2003). A *Markov Chain Monte Carlo*

(MCMC) method accomplishes this aim by generating a Markov Chain (see appendix B.7) whose chain elements are samples from the target distribution, and whose equilibrium distribution *is* the target distribution.

There are several ways to implement an MCMC method - one of them is the Metropolis-Hastings algorithm, which we will review here.

Since the Metropolis-Hastings algorithm is an example of an MCMC method, and thus an example of a Markov Chain, we expect it to have the property that each new sample that is drawn from the distribution depends only on the previous sample. This is the case exactly: Suppose we have some current sample x^t. We define a *proposal density* $Q(x)$ which depends on the current state: $Q(x) = Q(x; x^t)$. We then draw a sample from this distribution, x^p, and evaluate

$$\alpha = \frac{P(x^p)Q(x^t; x^p)}{P(x^t)Q(x^p; x^t)}. \tag{9.2}$$

The proposed value x^p is then accepted as a sample with probability α (or it is simply accepted if $\alpha > 1$). If x^p is accepted, we set $x^{t+1} = x^p$ - otherwise, we set $x^{t+1} = x^t$. The chain of samples x^t can then be shown to have $P(x)$ as its equilibrium distribution when $t \to \infty$ (McKay, 2003), and the chain is sampling from this distribution. With enough samples, the sample distribution approaches $P(x)$.

The proposal density Q is often set to be the same function for every point - a Gaussian with some fixed standard deviation, for instance - in which case the ratio $Q(x^t; x^p)/Q(x^p; x^t)$ just evaluates to one. In general, though, the shape of the proposal density may depend on the point in question, and it can be chosen such that the chain converges faster to its equilibrium position.

There is a period before the MCMC method begins sampling from the actual distribution: The chain might begin in a position in x space where it would not otherwise have ended up. This can be taken into account by disregarding a certain number of initial samples - the *burn-in* period.

There are some advantages of the MCMC methods over, say, grid-based evaluation, where one grids x space and directly calculates the PDF at each grid point: First, the mean values, confidence intervals, and so on may be estimated after a comparatively short time using MCMC methods, while the whole distribution function must be known in advance using the grid-based approach. The marginalisation of the PDF over some dimensions also becomes easier, as one only needs to disregard variations in those directions one wishes to marginalise over, and plot the sample distribution over the remaining directions. Further, the computational cost of the grid-based approach grows exponentially with the dimensionality of x space, while MCMC

methods don't (Christensen et al., 2001), so for higher-dimensional PDFs, the MCMC methods are the most resource-effective ones.

9.1.3 Cosmological applications and CosmoMC

Using the theory of Bayesian inference and the MCMC samplers, we may try to apply the above to the field of cosmology. We are then interested in answering the question: What cosmological model fits the data that we have in the best possible way? In terms of Bayesian concepts, we want to know the posterior distribution - the probability that some model is correct, given the data we have. This probability, then, according to Bayes' theorem, is proportional to the likelihood times the prior distribution. If we put reasonable physical limits on the parameters, but otherwise regard any value as equally possible, we assign to them a uniform prior. For acceptable parameter values, the prior then becomes parameter-independent. This is the situation with which we will mostly be working from here on. The likelihood function, then, is what really must be calculated in order to find the posterior distribution - but in what way can we quantify the probability of getting some set of data given some cosmological model?

The answer lies in the C_l's. By calculating the true C_l's for some model, using $CAMB$ or a similar program, and estimating the noise values (cosmic variance and instrumental noise) for the data we use, we can find the probability distribution function (which depends on the true C_l's and the noise estimates) for that model. Then, we insert the observed C_l's and determine the probability of observing those C_l's, given the model we are assuming. This then is exactly the likelihood function we are looking for, and all that is left to do is to map its properties in parameter space.

The CosmoMC software (Lewis & Bridle, 2002) is designed for this purpose. Using the MCMC method, it samples the likelihood function by evaluating the likelihood at each point, and using this as the $P(x)$ function in eq. (9.2). As we saw, the samples drawn this way are samples from the actual likelihood function, and its various properties may be calculated.

CosmoMC features

We will here mention some of the features of CosmoMC, as described in its `Readme` file (Lewis, 2010). We will focus on the features that will be explicitly used in the further analysis.

CosmoMC uses a fairly standard parametrisation of cosmological models, including parameters that are outside the scope of this text. The parameters encountered in chapter 6, for instance, are all found in CosmoMC. The tensor-to-scalar ratio r is also used as a parametrisation of the tensor contributions to the CMB.

One may use a variety of data sets and additional parameter constraints (effectively replacing the uniform prior with something else) with CosmoMC, and though we will here focus mostly on the $WMAP$ 5-year and 7-year data sets, and uniform priors, we will use some extra data sets and priors for an initial data run, as explained below.

The software has support for MPI parallelisation, meaning that one can effectively do several chains at once for the same distribution, making the sampling process quicker. The optimal number of chains to run is about eight. The usage of parallelisation also allows convergence testing. CosmoMC allows for two types of convergence tests. First, there is the *Gelman-Rubin* statistic R, which is the variance of the chain means divided by the mean of the chain variances, for each parameter. This value should be as close to unity as possible. Second, one may compare the confidence limits of each parameter across chains, to see if they converge.

CosmoMC also comes with a plot generator, which generates both two-dimensional marginalised contour plots, and one-dimensional marginalised distribution plots. In both cases, both the sample distribution and the mean likelihood at each point are shown (see e.g. figs. 9.1 and 9.2). This last detail allows for more information about the full distribution than would otherwise have been known by just ignoring variations in other parameters: The mean likelihood value at a point in marginalised parameter space tells us how good a fit we would expect if we drew a random model with that (or those) parameter values (Lewis & Bridle, 2002), while, as more thoroughly discussed in the next section, the marginalised distribution tells us what the probability of a certain parameter value is. For Gaussian distributions, the two will coincide, but for skew distributions in general they will not.

9.1.4 Statistics

A legitimate question to be asked when searching for the best cosmological model is, given some posterior distribution, which combination of parameters is the best? Is it the combination that maximises the likelihood function in n-dimensional parameter space (n being the total number of parameters sampled over), or is it perhaps the mean values of the one-dimensional marginalised distributions, or is it something completely different?

The answer to such questions depends on what we are asking. If we are asking for the model that gives the best fit to data, then the answer is the point in full-dimensional space which maximises the likelihood. If we are, on the other hand, attempting to answer the question, "Which value of parameter y is the most probable, given the data (and the underlying assumptions)?", the answer would be found in the marginalised one-dimensional distribution of y. The mean value of this distribution, however, is not what we are after - we want the most probable value of y. This value is the *peak* of y's marginalised distribution, which, if the posterior distribution is Gaussian, coincides with the mean value; if, on the other hand, the posterior is skew-symmetric, these values will differ.

The peak of a parameter's one-dimensional distribution will depend a good deal on the size of the bins into which we divide our parameter range, and is thus not a very reliable quantity. In order to say something a bit more informative and reliable, we may use the *Minimum Credible Interval* (MCI). This is the interval $[y_{min}, y_{max}]$ that contains a given percentage of the distribution, while also being the smallest such interval (that is, the distance between y_{min} and y_{max} is minimal). This interval will contain the one-dimensional peak point, and it represents a certain percentage confidence we may have that the true value of y lies within that interval (Hamann et al., 2007).

Another useful quantity is the actual likelihood value at a given point. When doing two different runs, using the exact same data but perhaps varying different parameters, one may compare the likelihood values at the ML points of both runs in order to find which of the models fits best. This method then allows a cross-run comparison of ML points, as long as the same data sets are being used.

CosmoMC has support for finding the full-dimensional ML point as well as the one-dimensional peaks and MCI limits, and it provides the likelihood values at the full-dimensional ML points.

9.2 Article reviews

We here review the articles mentioned in the introduction to this chapter. They draw on some of the theory above, but, as already mentioned, they are not based on the CosmoMC software.

We must first make clear that in order to parametrise the tensor contribution to the CMB, the authors do not use the r parameter introduced in this text,

but rather the tensor-to-scalar *quadrupole* ratio, denoted by R. These can be related using eqs. (17), (18) and (22) in Turner & White (1996) and eq. (4.33) in the current text to yield $r \approx 2R$, which the authors of the papers currently in question also use and cite. Throughout, they also use the relation $n_t = n_s - 1$, which follows if the scalar and tensor perturbations were generated by the same mechanism (which they were, in our treatment), and if there is no running of the scalar spectral index (which the authors assume). The pivot wavenumber is, as before, $k_p = 0.002$ Mpc^{-1}.

9.2.1 Zhao et al. (2009a)

This article is an analysis of the $WMAP$ 5-year data in search of primordial gravitational waves, and a forecast of the prospects of detecting tensor contributions to the CMB in the *Planck* mission. The authors first discuss the estimators of the autocorrelation spectra (i.e. BB, II, etc.) and the cross-correlation spectra, and their probability density functions (given some cosmological model). Using the IE probability density function to perform certain hypothesis tests and a likelihood function analysis for R, the authors claim that this spectrum contains a 'hint' of tensor contributions, and that the most likely value of R is 0.240. The paper also contains a comparison between the usage of the IE spectrum vs. the BB spectrum, and concludes that for realistic cases, and for large enough tensor contributions, the IE mode will be a better detector of primordial gravitational waves.

Comments

This article is not our main focus in this text. We may make a few points before moving on to the main paper, however:

The probability density functions for the estimators must take into account the noise values for each l. The noise values used by the authors are illustrated in their fig. 6, showing that they approximate the true noise (which is correlated between different l values) by uncorrelated noise. This could well be a good approximation, but we nonetheless note the point here for future reference.

Another point is that throughout the analysis, the authors limit themselves to the range $l = 2 - 100$, for the reason that gravitational waves only affect this range. We shall comment more on this below, after reviewing the second article.

9.2.2 Zhao et al. (2009b)

In this article, the authors follow up their previous article with a more thorough likelihood analysis of the parameter R, this time taking into account the fact that R is degenerate with respect to other parameters. The way they accomplish this is to employ a self-made MCMC sampler, varying the parameters R, n_s and A_s. (This implicitly also means varying n_t since $n_t = n_s - 1$). Fixing all other parameters using their mean values from the $WMAP$ 5-year analysis (Komatsu et al., 2009) (last column of Table 1), they perform the MCMC analysis using 10,000 samples, with the probability density functions and noise values being the same as for the previous paper, and using the II and IE $WMAP$ 5-year data points for $l = 2 - 100$. From this analysis, they obtain the maximum likelihood values

$$R = 0.229, \quad n_s = 1.086, \quad A_s = 1.920 \times 10^{-9}, \tag{9.3}$$

which, when using $r \approx 2R$, goes against the results obtained by the $WMAP$ team using the 5-year data - namely, $r < 0.22$. Marginalising the three-dimensional distribution to a one-dimensional distribution for each parameter, they obtain the following one-dimensional peak values and their 68.3% MCI limits:

$$R = 0.266 \pm 0.171, \quad n_s = 1.107^{+0.087}_{-0.070}, \quad A_s = (1.768^{+0.307}_{-0.245}) \times 10^{-9}. \tag{9.4}$$

This means that this is a detection of a nonzero R on a 2σ level - not a solid detection by any standards (not even cosmological ones). However, as in the previous paper, this 'hints' of tensors being present in the CMB.

The authors have only used a limited l range in their analysis, and they now want to show that using a more extended range can give the wrong results, and why this is so. They perform the same MCMC analysis on the range $l = 101 - 220$, and find that this changes the ML results considerably:

$$R = 0.022, \quad n_s = 0.923, \quad A_s = 2.65 \times 10^{-9}. \tag{9.5}$$

This R value has a large uncertainty, as their fig. 5. shows. This, the authors say, is due to the gravitational waves not having much effect on $l > 100$. What is more interesting is the low n_s value. For this range, one gets a red spectrum , while for the low l range, one gets a blue spectrum. Further, when marginalising the full distribution for the high l range, one finds that the two values of n_s do not overlap inside a 68.3% confidence interval. This shows that assuming n_s to be constant for all l is an erroneous assumption, and that this is the cause of discrepancy between their results and the $WMAP$ 5-year results. To show this more explicitly, they run the analysis again, but this time over the range $l = 2 - 220$. They end up with an R ML value

similar to eq. (9.5), and an n_s ML value that lies between those of eqs. (9.3) and (9.5).

The rest of the paper is another forecast for the *Planck* mission, using the maximum likelihood results obtained so far. This does not interest us here - we focus on the likelihood analysis. The authors conclude that the assumption of an l-independent (or running) spectral index gives incorrect results, and one should only consider the l range that is affected by gravitational waves when searching for these.

Comments

We will here comment shortly on the above likelihood analysis and its interpretation. First, we must note that it is not correct to say that we should only use the l range affected by tensors when looking for these. On the contrary, since the low l range is so degenerate (as we saw in chapter 6), we should use as large an l range as possible, since this will help constrain other parameters, which again helps constrain the tensor contribution.

Of course, this assumes that we are using the correct parametrisation, which the above analysis might seem to suggest that we are not. There is a tension between the $l = 2-100$ results and the $l = 101-220$ results that may indicate that the assumption of a scale-independent spectral index is wrong. They do not overlap at a 1σ-level, and though a 2σ detection is not very robust, this may hint that some of our assumptions may actually be false: If the spectral index is indeed constant, we should find that all data sets used are consistent with some constant value of n_s, and if we view the $l = 2 - 100$ range and the $l = 101 - 220$ range as two different data sets, we have found two sets that are (mildly) inconsistent with the hypothesis of a constant spectral index.

This also holds for R - there was a $> 1\sigma$ detection of a nonzero R for $l = 2 - 100$. In this case, the uncertainties for $l = 101 - 220$ were so large that a nonzero R could easily be incorporated - but in order for $R = 0$ to be incorporated into the $l = 2 - 100$ analysis, we would have to go beyond the 1σ level. It may seem, then, that there again are 'hints' of a nonzero R in the CMB.

However, to infer from this that the (wrongly) assumed constancy of n_s leads to a (wrong) model where the tensor contribution is negligible, is to go beyond the evidence, as we do not know what the solution to the n_s-inconsistency might be, and this solution might still yield a negligible R. This claim will be tested further presently.

The final point to make here, and to keep in mind, is that the $WMAP$ team uses the consistency relation (4.33) in their analyses. In light of the whole controversy we have been studying so far, this seems odd: We want to test the tensor contribution to the CMB in order to test whether the theory of inflation gives the correct predictions, but we do that partly by assuming the correctness of the inflationary predictions (the consistency relation)! The $WMAP$ team states that they use this relation in order to reduce the number of free parameters. Below, we will among other things do a likelihood analysis of the $WMAP$ 7-year data and other data sets, using the relation $n_t = n_s - 1$ instead of the consistency relation, which is the relation Zhao et al. also use.

9.3 Improving the likelihood analysis

The likelihood analysis carried out in Zhao et al. (2009b) can be improved in some respects.

First, one may utilise the CosmoMC software described above. This has several advantages: It takes into account the complete noise covariance matrix, so that one needs not use the uncorrelated noise approximation. Further, the software has been around for many years and bugs and errors has thus been weeded out in a more complete way than that of a 'home-made' MCMC sampler as the one used above. It also comes with plot generators so that plotting of likelihood contours etc. becomes much easier. Finally, it allows utilisation of the EE and BB data points in addition to the IE and II ones[1], and it allows simulation of the lensing effects on the CMB. The obvious disadvantage of using this software is the implementation of non-standard parameters (such as R), which means familiarising oneself with code written by someone else and modifying it according to one's needs. This is a significant obstacle, but once it has been traversed, one may have more confidence in the results of the sampler than if one made the program oneself.

Second, there are more parameters affecting the power spectrum at the l ranges we are considering than the ones investigated in Zhao et al. (2009b). Most importantly, we saw in chapter 6 that the parameter τ can affect the power spectrum in such a way that by adjusting τ and A_s, one may produce a no-tensor power spectrum that is very close to a spectrum where tensors are present. Thus, by including τ in the likelihood analysis, one may end up with different results than those above.

[1] Although only the $l = 2 - 23$ data points are used for the EE and BB spectra.

Third, one should also be aware of the limitations of using the best-fit values from the $WMAP$ analyses: These values come from an analysis where it is assumed that $r = 0$. This means that the best-fit values we get are biased, as they assume no tensor contribution from the outset. This may influence the results of our analyses. Unfortunately, even though the $WMAP$ team also carries out a full-fledged analysis with all parameters, allowing r to vary as well, they do not provide any values from this analysis except the limits on r, which are not very useful in our discussion (in addition, we have mentioned that the $WMAP$ team uses the consistency relation in this more complete analysis, which would render their results useless for us anyway).

Because of this, a better approach would be to do an initial analysis which takes into account most available data and prior constraints, allowing for r to be nonzero, and varying an extended set of cosmological parameters. Then the relation $n_t = n_s - 1$ can also be used, which is assumed by the authors of Zhao et al. (2009b). The best-fit values from this analysis can subsequently be used to fix parameters in analyses such as the one we have reviewed.

Again, we are faced with the question of what 'best-fit values' actually mean. We here argue that it makes most sense to use the full-dimensional ML values as the best-fit values: These are taken from the model which best fits the data. This then contrasts with what was done in Zhao et al. (2009b); there, the authors used the mean values of the one-dimensional parameter distributions. As long as the distributions are Gaussian, this means that they used the most probable value of each parameter. There is no clear-cut answer to which of these approaches is the correct one, but the method we advocate has the benefit of using the parameter values that most likely describe our Universe.

9.3.1 Caveat

It is important to be aware of the assumptions one makes when doing a likelihood analysis. Comparing two analyses with different underlying assumptions may give conflicting results. For example, say we do two analyses (using the same or different data sets): One where we vary n_s, A_s, and R, and another where τ is varied in addition to these. The parameters we do not vary must be fixed to some best-fit value. However, these analyses are not really answering the same question: The first asks, "If we vary these three parameters, assuming τ (and all other parameters) to have a certain value, what will be the ML points and one-dimensional distributions for those three parameters?", while the second does not assume a certain

value of τ. Asking two different questions will in general give two different answers, so we should not uncritically compare the two analyses.

9.4 Repeated analysis using the improved method

In this section, we will utilise the improvements outlined above and repeat the analysis in Zhao et al. (2009b). The modifications made to CosmoMC in order to implement R instead of r are described in appendix C.1. In addition, we here describe the relevant input parameters and the exact procedure, to make the results as replicable as possible.

9.4.1 Input parameters and data used

We use the modified version of CosmoMC above to carry out a likelihood analysis similar to the one in Zhao et al. (2009b).

As the work behind this text was being carried out, $WMAP$ released their 7-year data, and CosmoMC was updated to include this shortly after. We have therefore opted to do two separate analyses: One using the 5-year $WMAP$ data, and one using the 7-year data. We do this because the 5-year data is what the authors of Zhao et al. (2009b) used, while the 7-year data should improve the results with better parameter constraints. For the repeated analysis, we only use the $WMAP$ data sets, as these are what the original likelihood analysis was based on. They are incorporated through the $WMAP$ likelihood software, which can be found at the $LAMBDA$ website (NASA, 2010).

Based on our discussion about what values at which to fix the parameters we do not vary, we choose the solution mentioned there: We do an initial analysis, described more fully below, where we use more types of data sets, allow gravitational waves and use the relation $n_t = n_s - 1$. This will give a certain best-fit model (the full-dimensional ML values), and for the rest of the analyses, we use this model to fix the parameters that we do not vary.

We always run eight chains, and we check all parameters for convergence, using the Gelman-Rubin statistic, demanding that $R-1 < 0.03$. In addition, we also always use the CosmoMC feature to check for convergence of confidence limits between chains, setting the convergence error equal to 0.2, and set the fraction of the distribution tail to check equal to 0.025 (these were the default values). The number of samples used in the current analysis will

therefore vary from run to run, which again differs from the original likelihood analysis, where 10,000 samples were used, regardless of convergence. Further, we cut away burn-in samples using the burn-in estimates provided by the CosmoMC software, at least doubling the estimate provided just to be sure.

For two-and-one-dimensional marginalisation, we use 30 bins along each parameter axis. This number was chosen by trial and error as a number that gave decent numerical accuracy while maintaining a certain smoothness of the distributions. We also use the option in CosmoMC to smooth each bin with a Gaussian kernel.

We do seven different runs for both data sets:

- Run 1: $l = 2 - 100$, varying n_s, A_s, and R
- Run 2: $l = 101 - 220$, varying n_s, A_s, and R
- Run 3: $l = 2 - 220$, varying n_s, A_s, and R
- Run 4: $l = 2 - 220$, varying n_s and A_s
- Run 5: $l = 2 - 100$, varying τ, n_s, A_s, and R
- Run 6: $l = 2 - 220$, varying τ, n_s, A_s, and R
- Run 7: $l = 2 - 220$, varying τ, n_s, and A_s

The first three of these are essentially a duplication of the original likelihood analysis, except for the changes associated with using CosmoMC. Runs 5 and 6 include τ in the analysis, cf. the above discussion. Runs 4 and 7 use the (full-dimensional) ML value for R obtained from the $l = 2 - 100$ analysis (the top run). We do not include τ in the $l = 101 - 220$ analysis since this l range contains too little data to constrain τ in addition to the three others (this statement has been verified by trial). This means that direct comparison between, say, runs 2 and 5, is not strictly correct for the reason we argued in sec. 9.2.2: In one of them, we fix τ, while in the other, we let it vary.

As mentioned above, in addition to the above runs, we do an initial analysis: We run CosmoMC with more types of data sets and some priors, i.e., CMB data (*WMAP* 7-year data (NASA, 2010)), matter power spectrum data (the Sloan Digitial Sky Survey (SDSS), fourth data release (Adelman-McCarthy et al., 2006)), supernova observations (SDSS, ESSENCE (Miknaitis et al., 2007), SNLS (Balland et al., 2009), HST (Riess et al., 2009), and various

low-redshift supernova observations), Lyman-alpha data (SDSS), and priors on the age of the Universe, the Hubble parameter (from the HST), and Ω_b (from the theory of Big Bang Nucleosynthesis). In addition, measurements of baryonic acoustic oscillations (BAO) are incorporated, which provides consistency checks between redshifts and physical distances. The l range for the $WMAP$ data is $l = 2 - l_{max}$, where l_{max} denotes the maximum value of l allowed by the likelihood software. This value varies by type of power spectrum: $l^{II}_{max} = 1200$ and $l^{IE}_{max} = 800$ (as noted above, the EE and BB spectra have an l_{max} of 23, regardless).

In the initial run, we vary seven parameters: $\Omega_b h^2$, $\Omega_{DM} h^2$, θ, τ, n_s, A_s, and r, where Ω_{DM} is the density parameter of dark matter, $h = H_0/100$, and θ is the ratio of the sound horizon to the angular diameter distance (see eq. (9.7). Any further details of this parameter is prohibited by the scope of this text, but it is used instead of H_0 since it is less correlated with other parameters). Varying these parameters means that we have made certain assumptions: We assume that the equation of state of the 'dark energy' Λ is $w = -1$, that neutrinos are massless, and that our Universe is flat. We also use the relation $n_t = n_s - 1$, and assume no running of the spectral index.

Doing an initial run as described above serves two purposes: First, as mentioned above, it gives better values for the parameters we intend to fix later than the $WMAP$ analysis (Komatsu et al., 2010) does, since our analysis does not rule out gravitational waves to begin with, and it uses the above-mentioned relation between the scalar and tensor spectral indices, which the authors of Zhao et al. (2009b) also use, but which the $WMAP$ team doesn't. Second, we will then obtain results to which we can compare our later results, though, based on the discussion in sec. 9.3.1, direct comparisons are out of place.

9.4.2 Results

The results of the initial run are shown in table 9.1, where we list the full-dimensional ML point and the MCI limits for each variable, cf. the discussion in section 9.1.4. The results of the subsequent runs are shown in table 9.2.

We do not plot all the marginalised one-and-two-dimensional sample distributions, since it would take up far too much space, so we focus on plotting the most interesting features. In fig. 9.1, we plot the one-dimensional sample distribution for n_s for runs 1, 2, and 5, both for 5-year and 7-year data, since the incompatibility of these distributions was a central point in Zhao et al. (2009b) (as an aside, from these figures is evident the phenomenon

Parameter	ML value	One-dimensional peaks and MCI limits
		$X^{\substack{68\%\uparrow,\ 95\%\uparrow \\ 68\%\downarrow,\ 95\%\downarrow}}$
$\Omega_b h^2$	0.0226	$0.0228^{\substack{0.0233,\ 0.0238 \\ 0.0223,\ 0.0218}}$
$\Omega_{DM} h^2$	0.1177	$0.1176^{\substack{0.1203,\ 0.1229 \\ 0.1145,\ 0.1118}}$
θ	1.040	$1.040^{\substack{1.043,\ 1.045 \\ 1.038,\ 1.036}}$
τ	0.0810	$0.0876^{\substack{0.1028,\ 0.1176 \\ 0.0749,\ 0.0625}}$
n_s	0.963	$0.969^{\substack{0.983,\ 0.996 \\ 0.957,\ 0.945}}$
$\log(10^{10} A_s)$	3.207	$3.203^{\substack{3.243,\ 3.281 \\ 3.165,\ 3.123}}$
r	0.015	$0.000^{\substack{0.076,\ 0.168 \\ 0.000,\ 0.000}}$

Table 9.1: ML points and marginalised one-dimensional peaks and MCI limits for a likelihood analysis for seven parameters, taking into account CMB, power spectrum, and supernova data, in addition to various prior constraints on some parameters, described in the text. MCI limits are given with both 68% and 95% confidence.

mentioned earlier: The sample distribution does not perfectly coincide with the mean likelihood distribution). We also plot the two-dimensional sample distributions for all parameters from the 5-year version of run 5 (fig. 9.2), to illustrate the parameter degeneracy of this l range, which we also encountered in chapter 6.

9.4.3 Discussion

We have expanded the original analysis (Zhao et al., 2009b) in several ways: Using the proper noise data, including the BB and EE modes, including τ as a variable parameter, using better fixed parameter values, and using the 7-year $WMAP$ data in addition to the 5-year data. Despite this, our findings are not very different from those of the original analysis. Below, we discuss the general features for each parameter.

First, we markedly see the degeneracies of the various parameters in fig. 9.2, which is a contour plot from run 5, especially between the parameters A_s, n_s, and R.

For the runs that include the parameter τ, we see that using only the $l = 2 - 100$ range raises its value compared to more complete data sets, but all ML and peak values are well within the MCI limits, which are quite large. Including it in the analysis does not have a very large effect, except that it raises the value of R somewhat, especially in the $l = 2 - 100$ runs. This

Parameter	ML value		One-dimensional peaks and MCI limits $X^{\substack{68\%\uparrow,\ 95\%\uparrow \\ 68\%\downarrow,\ 95\%\downarrow}}$	
	5-year	7-year	5-year	7-year
Run 1 ($l = 2 - 100$, varying n_s, A_s, and R)				
n_s	1.038	1.036	$1.047^{\substack{1.126,\ 1.201 \\ 0.994,\ 0.949}}$	$1.049^{\substack{1.112,\ 1.186 \\ 0.996,\ 0.949}}$
$\log(10^{10}A_s)$	3.066	3.067	$3.047^{\substack{3.149,\ 3.220 \\ 2.896,\ 2.738}}$	$3.050^{\substack{3.145,\ 3.217 \\ 2.916,\ 2.775}}$
R	0.104	0.107	$0.069^{\substack{0.251,\ 0.466 \\ 0.000,\ 0.000}}$	$0.138^{\substack{0.229,\ 0.423 \\ 0.000,\ 0.000}}$
Run 2 ($l = 101 - 220$, varying n_s, A_s, and R)				
n_s	0.925	0.953	$0.961^{\substack{1.031,\ 1.104 \\ 0.893,\ 0.828}}$	$0.979^{\substack{1.062,\ 1.138 \\ 0.914,\ 0.846}}$
$\log(10^{10}A_s)$	3.291	3.234	$3.203^{\substack{3.355,\ 3.489 \\ 3.066,\ 2.909}}$	$3.164^{\substack{3.316,\ 3.455 \\ 3.005,\ 2.844}}$
R	0.028	0.245	$0.335^{\substack{1.044,\ 2.050 \\ 0.000,\ 0.000}}$	$0.112^{\substack{1.144,\ 2.161 \\ 0.000,\ 0.000}}$
Run 3 ($l = 2 - 220$, varying n_s, A_s, and R)				
n_s	0.976	0.991	$0.998^{\substack{1.034,\ 1.073 \\ 0.973,\ 0.949}}$	$1.010^{\substack{1.048,\ 1.091 \\ 0.980,\ 0.955}}$
$\log(10^{10}A_s)$	3.185	3.158	$3.136^{\substack{3.191,\ 3.237 \\ 3.068,\ 2.984}}$	$3.116^{\substack{3.179,\ 3.229 \\ 3.042,\ 2.951}}$
R	0.001	0.021	$0.000^{\substack{0.100,\ 0.210 \\ 0.000,\ 0.000}}$	$0.000^{\substack{0.121,\ 0.243 \\ 0.000,\ 0.000}}$
Run 4 ($l = 2 - 220$, varying n_s and A_s)				
n_s	1.018	1.024	$1.014^{\substack{1.039,\ 1.056 \\ 0.999,\ 0.980}}$	$1.021^{\substack{1.044,\ 1.065 \\ 1.004,\ 0.984}}$
$\log(10^{10}A_s)$	3.098	3.089	$3.098^{\substack{3.134,\ 3.172 \\ 3.057,\ 3.026}}$	$3.092^{\substack{3.127,\ 3.164 \\ 3.050,\ 3.009}}$
Run 5 ($l = 2 - 100$, varying τ, n_s, A_s, and R)				
τ	0.0975	0.0941	$0.1021^{\substack{0.1217,\ 0.1454 \\ 0.0798,\ 0.0624}}$	$0.0937^{\substack{0.1128,\ 0.1312 \\ 0.0791,\ 0.0638}}$
n_s	1.076	1.063	$1.095^{\substack{1.174,\ 1.247 \\ 1.028,\ 0.969}}$	$1.084^{\substack{1.145,\ 1.221 \\ 1.016,\ 0.967}}$
$\log(10^{10}A_s)$	3.025	3.041	$2.982^{\substack{3.122,\ 3.209 \\ 2.862,\ 2.717}}$	$3.028^{\substack{3.123,\ 3.212 \\ 2.892,\ 2.760}}$
R	0.167	0.146	$0.208^{\substack{0.341,\ 0.520 \\ 0.044,\ 0.000}}$	$0.136^{\substack{0.293,\ 0.447 \\ 0.038,\ 0.000}}$
Run 6 ($l = 2 - 220$, varying τ, n_s, A_s, and R)				
τ	0.0876	0.0876	$0.0869^{\substack{0.1079,\ 0.1276 \\ 0.0723,\ 0.0573}}$	$0.0879^{\substack{0.1056,\ 0.1227 \\ 0.0747,\ 0.0616}}$
n_s	0.980	0.993	$1.008^{\substack{1.044,\ 1.088 \\ 0.977,\ 0.952}}$	$1.020^{\substack{1.056,\ 1.102 \\ 0.985,\ 0.959}}$
$\log(10^{10}A_s)$	3.191	3.166	$3.139^{\substack{3.199,\ 3.245 \\ 3.068,\ 2.982}}$	$3.134^{\substack{3.186,\ 3.235 \\ 3.047,\ 2.955}}$
R	0.002	0.022	$0.000^{\substack{0.109,\ 0.225 \\ 0.000,\ 0.000}}$	$0.015^{\substack{0.128,\ 0.249 \\ 0.000,\ 0.000}}$
Run 7 ($l = 2 - 220$, varying τ, n_s, and A_s)				
τ	0.0882	0.0912	$0.0942^{\substack{0.1079,\ 0.1290 \\ 0.0709,\ 0.0568}}$	$0.0927^{\substack{0.1058,\ 0.1206 \\ 0.0759,\ 0.0622}}$
n_s	1.020	1.028	$1.021^{\substack{1.045,\ 1.066 \\ 1.003,\ 0.981}}$	$1.030^{\substack{1.051,\ 1.071 \\ 1.008,\ 0.988}}$
$\log(10^{10}A_s)$	3.108	3.101	$3.101^{\substack{3.149,\ 3.191 \\ 3.065,\ 3.026}}$	$3.094^{\substack{3.137,\ 3.176 \\ 3.057,\ 3.017}}$

Table 9.2: The unmarginalised ML points and one-dimensional peaks and MCI limits resulting from a likelihood analysis using both 5-year and 7-year $WMAP$ data for various l ranges, and with various parameters. MCI limits are given with both 68% and 95% confidence.

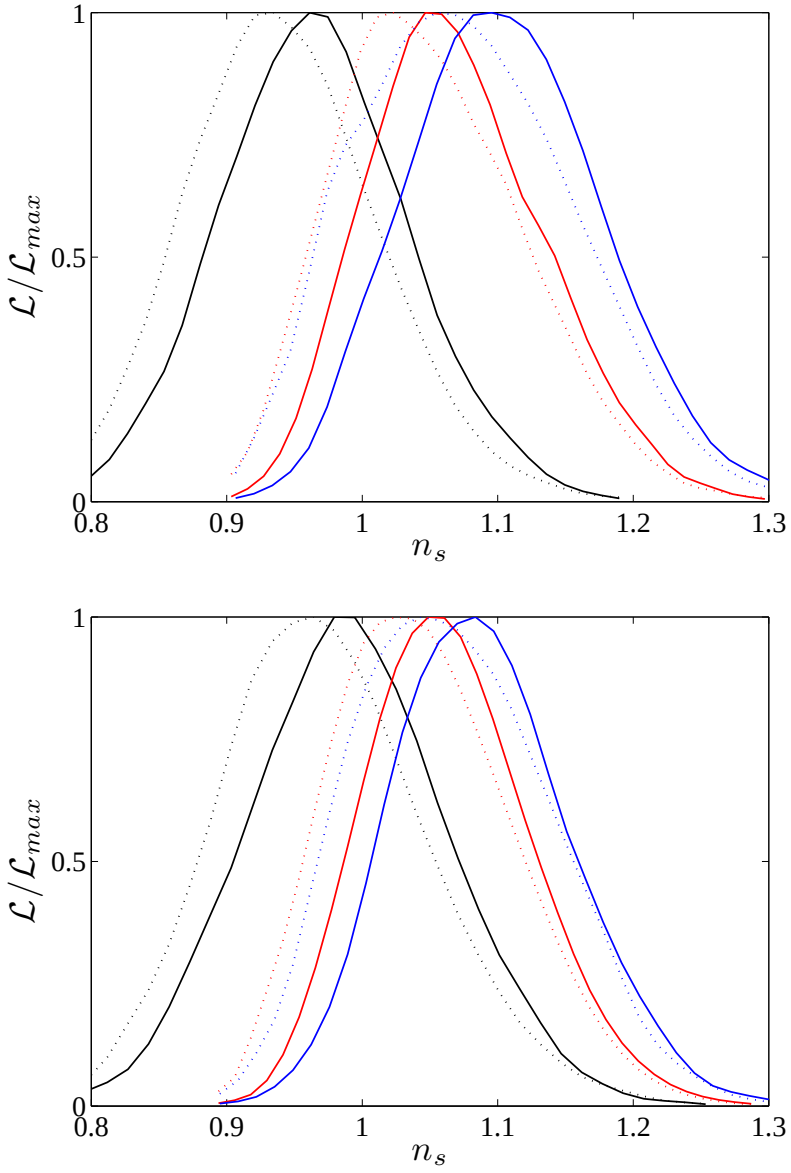

Figure 9.1: Marginalised sample distributions for n_s for three different runs: The $l = 101 - 220$ run (black), the $l = 2 - 100$ run where τ is not allowed to vary (red), and the $l = 2 - 100$ run where τ is allowed to vary (blue). Top picture is for $WMAP$ 5-year data, while bottom picture is for the 7-year data. Both sample distributions (solid lines) and mean likelihood values (dashed lines) are shown.

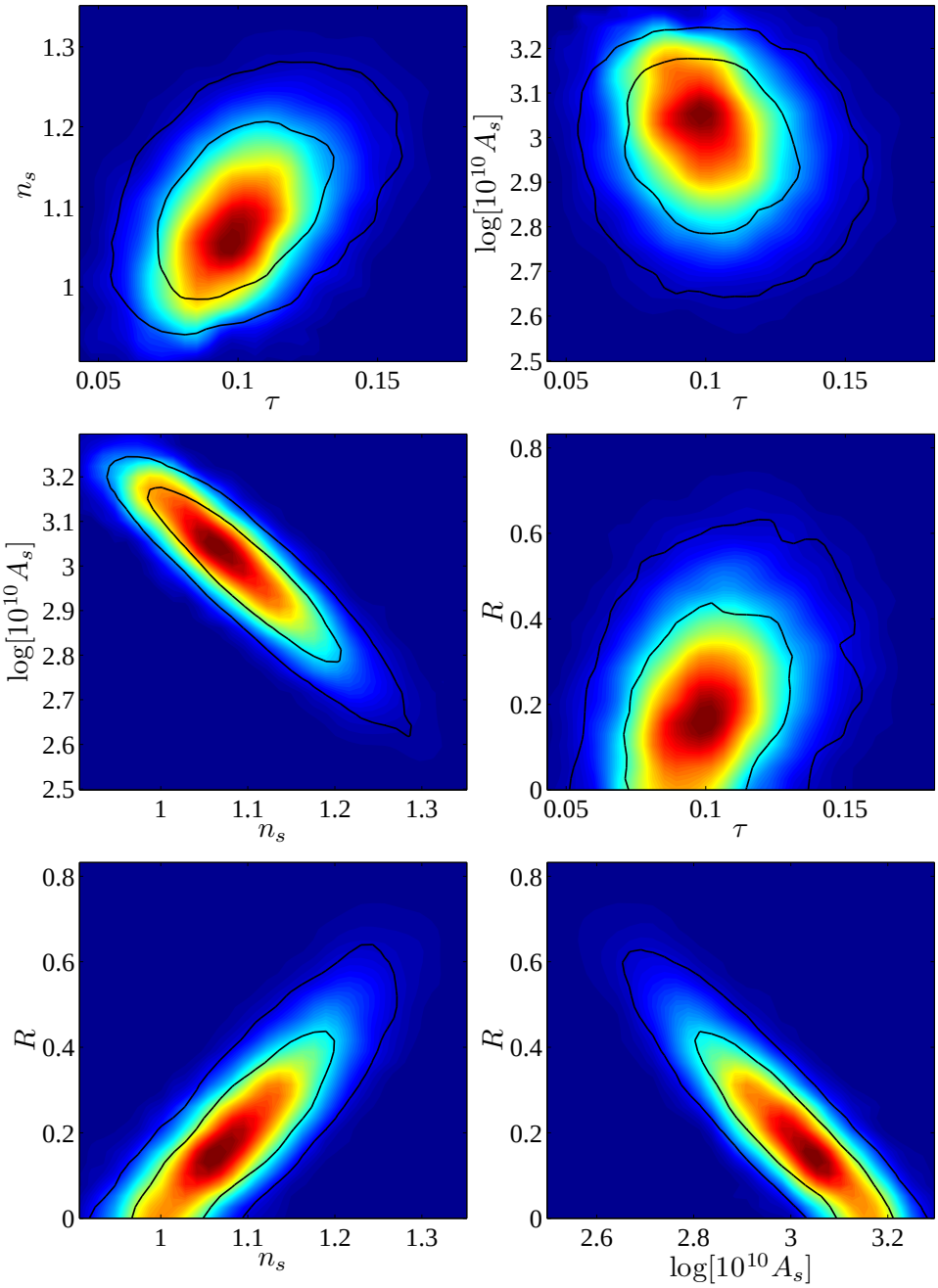

Figure 9.2: Two-dimensional marginalised sample distributions for the parameters τ, n_s, A_s, and R, based on the range $l = 2-100$, and using $WMAP$ 5-year data. The solid lines are 68% and 95% contour lines of the sample distribution, while the colors indicate the mean likelihood at that point.

indicates that even though the fixed best-fit value of τ is rather low, it is more likely that other parameters than R, like n_s, compensate for this.

The parameter A_s vary, as expected, with the other parameters, but there are no particularly interesting results concerning this parameter.

The parameter n_s, however, is more interesting. There are two points to note here: First, we find, in contrast to the original analysis, no discrepancy between the run 1 MCI limits and the run 2 MCI limits - they overlap within a 68% interval - though only barely. This is also shown in fig. 9.1, where the one-dimensional distributions for n_s are shown for the various runs and for the two data sets.

Second, we see a small discrepancy between runs 5 and 7, and the initial run: The 68% MCI limits for n_s do not overlap here. However, this result should be treated with caution, since, as we have noted several times already, in runs 5 and 7 we are fixing parameters that we are not in the initial run.

For the parameter R, we basically find the same results as in Zhao et al. (2009b), except that our $l = 2-100$ values are consistently and substantially lower: Our run 1 gives a value of R which is roughly half the value found in the corresponding run in the original analysis. We also see that the ML value of R changes rather drastically in run 2 between the 5-year and 7-year runs, but we see that the uncertainties are very large for this run, which is a consequence of the low impact gravitational waves have on this l range. We find lower limits for R in run 5, again 'hinting' of a nonzero R, but of course, this only holds for the models where our fixed values of certain parameters are correct.

To conclude, we find that the analysis in Zhao et al. (2009b) is quite consistent with ours, except that we find a lower R and that in our analysis, the n_s distributions for runs 1 and 2 *do* overlap, albeit only barely. Thus, there is no longer any real conflict between these data sets.

9.5 Testing the l-dependency of n_s

In this section, we will investigate the claim in Zhao et al. (2009b) that the assumption of an l-independent n_s is incorrect and leads to the abandonment of a non-negligible R even though hints of such an R is present when we limit the l range. We do this by implementing a step-like n_s, so that for $l < 100$, we have a (constant) value of n_s, denoted n_s^{low} while for $l > 100$, we have another value of n_s, denoted n_s^{high}. Instead of working directly with

the l's, we use the fact that a given l can be associated with a k:

$$l \approx k d_A, \tag{9.6}$$

where d_A is the *angular diameter distance*. For flat universes, which we currently assume, this quantity is given by (Elgarøy et al., 2002)

$$d_A \approx \frac{2c}{H_0 \Omega_d^{0.4}}, \tag{9.7}$$

where Ω_d is the density parameter of dust. Using these relations, we can then translate an l-dependent step function into a k-dependent one. We choose to work with the k's for the reason that it becomes a bit easier to implement, and because a k-dependent n_s is slightly easier to justify physically than an l-dependent one (since the k's actually correspond to a physical scale, while the l's represent a decomposition of our sky - though they are closely related to physical scales, of course). However, we emphasise that there is no actual physical justification for using a step-function for n_s, except that we wish to test whether we can use such a function to retain an R value of ≈ 0.2 over a larger range of l's, as the authors of the original analysis claim we can. If their assessment is correct, we should expect the following from the use of a step-like n_s:

- The R ML values and MCI limits should be raised for the $l = 2 - 220$ range compared to the analysis in the previous section (runs 3 and 6).
- Over the range $l = 2-220$, we should find that n_s^{high} and n_s^{low} approach the values found in run 2, and runs 1 and 5 of the previous analysis, respectively. This should at least be the case when we fix R to have the value from the $l = 2 - 100$ analysis above.
- Conversely, if we fix n_s^{high} and n_s^{low} to have the corresponding values obtained in the previous section, we should get R-values comparable to those of the $l = 2 - 100$ analyses.

We choose to use the k value that corresponds to $l = 100$ as the 'step-position' since this is where the line has been drawn both in the original analysis and in our analysis so far. The necessary changes made to CosmoMC in order to implement such a spectral index can be found in appendix C.2.

Why not use running of the spectral index?

An obvious way to implement a l-dependent spectral index would simply be to use the running of the spectral index. The reason we don't do this

here is that in the paper Zhao et al. (2009b), the authors do include some comparisons with the $WMAP$ analysis that includes a running. They do not comment much on this, but they seem to say that this is not a suitable enough parametrisation.

Also, we want to engage with the claims made in that paper on the terms of that paper, not on our own terms, and a step-like n_s is the most direct way to do this, even though it may have limited physical basis. One might say that we are not directly testing a physical theory here; rather, we are testing a claim that, if correct, will need some physical justification at some point.

9.5.1 Input parameters and data used

We use the same data sets and ML parameters as described in section 9.4.1, and all CosmoMC parameters except the obvious ones are kept the same as in that section. We always use `l_step = 100`.

We do these runs for both data sets:

- Run 1: Varying $n_s^{low}, n_s^{high}, A_s$, and R.
- Run 2: Varying $\tau, n_s^{low}, n_s^{high}, A_s$, and R.
- Run 3: Varying $\tau, n_s^{low}, n_s^{high}$, and A_s
- Run 4: Varying τ, A_s, and R
- Run 5: Varying n_s^{low}, n_s^{high}, and A_s
- Run 6: Varying A_s and R

We do all runs for two l ranges: $l = 2 - 220$, and $l = 2 - l_{max}$. For 5-year data, the l_{max} values are a little different from before: $l_{max}^{II} = 1000$ and $l_{max}^{IE} = 450$. Again, the EE and BB ranges only go up to $l = 23$. When we are not varying R, we use its ML value from the $l = 2 - 100$ analysis of the previous section, and when not varying n_s^{low} and n_s^{high}, we use their ML values from the $l = 2 - 100$ and $l = 101 - 220$ analyses, respectively.

9.5.2 Results

We summarise the results of the analysis in tables 9.3 and 9.4. In addition to these tables, we show a comparison of the likelihood values of the ML models

for some of the runs, both from this section and from the previous analysis, in table 9.5 (these likelihood values say nothing by themselves; they only enable comparison of various models). We also show some contour plots from the 5-year $l = 2 - 220$ run 2, in order to again show the degeneracies of the newly introduced parameters to other parameters.

9.5.3 Discussion

The analysis of the incorporation of a l-dependent step-like n_s seems to have the following general features: First, using the the whole l range at our disposal instead of just the $l = 2 - 220$ range always lowers the ML values and MCI limits for R to a close to negligible value, and lowers the values of both n_s^{high} and n_s^{low}. Second, runs with a constant n_s yield almost the same R values as the corresponding ones (same l range) with a step-like n_s (compare run 3 and run 6 in table 9.2 with run 1 and 2 in table 9.3, respectively). Finally, if R is allowed to vary, we in general get a lower n_s^{high} than n_s^{low}, which we expected, but they are always closer together than what we would have expected from the previous analysis (e.g., runs 1, 2, and 5 in table 9.2). Also, the one-dimensional distributions for n_s^{low} and n_s^{high} always overlap within the 68% interval, so there is no reason for abandoning the constant-n_s models based on our results.

Fixing R does not give values for n_s^{high} and n_s^{low} that corresponds to the values obtained for the $l = 2 - 100$ range and the $l = 101 - 220$ range in the previous analysis. On the contrary, we see that fixing R actually drives the n_s's to a common value (runs 3 and 5 of tables 9.3 and 9.4), and we even see that n_s^{low} in some of the cases actually has a *lower* value than n_s^{high}, especially for the $l = 2 - l_{max}$ runs, which is in direct opposition to what we would expect. Likewise, fixing the n_s-values to be what we expect them to be from the previous analysis, we obtain (runs 4 and 6) negligible values for R - more negligible than they are without the fixation.

Finally, from table 9.5, we see that the ML models from the constant-n_s analysis fit only marginally worse than the ML models from the step-like n_s analysis, and that when fixing either R or n_s, one ends up with a model that fits worse than the constant-n_s analysis, whether one includes τ in the analysis or not.

All this taken together seems to imply that the expectations we had before the analysis have not been met, with the possible exception that n_s^{low} and n_s^{high} do indeed obtain different values, with a blue spectrum for the former and a red spectrum for the latter, when we only include the range $l = 2-220$, and don't fix R. However, they are not as blue or red as expected, and the

Parameter	ML value		One-dimensional peaks and MCI limits $X^{\substack{68\%\uparrow,\ 95\%\uparrow \\ 68\%\downarrow,\ 95\%\downarrow}}$	
	5-year	7-year	5-year	7-year
Run 1 (varying n_s^{low}, n_s^{high}, A_s, and R)				
n_s^{low}	0.993	1.007	$1.018^{1.077,\ 1.142}_{0.943,\ 0.881}$	$1.021^{1.084,\ 1.148}_{0.952,\ 0.887}$
n_s^{high}	0.983	0.987	$1.004^{1.037,\ 1.076}_{0.973,\ 0.949}$	$1.016^{1.049,\ 1.092}_{0.982,\ 0.955}$
$\log(10^{10}A_s)$	3.172	3.165	$3.124^{3.189,\ 3.234}_{3.060,\ 2.978}$	$3.117^{3.176,\ 3.229}_{3.039,\ 2.947}$
R	0.004	0.004	$0.000^{0.105,\ 0.231}_{0.000,\ 0.000}$	$0.015^{0.125,\ 0.262}_{0.000,\ 0.000}$
Run 2 (varying τ, n_s^{low}, n_s^{high}, A_s, and R)				
τ	0.0824	0.0926	$0.0874^{0.1090,\ 0.1286}_{0.0730,\ 0.0561}$	$0.0900^{0.1076,\ 0.1237}_{0.0762,\ 0.0627}$
n_s^{low}	1.004	1.027	$1.021^{1.093,\ 1.160}_{0.956,\ 0.888}$	$1.044^{1.111,\ 1.181}_{0.962,\ 0.896}$
n_s^{high}	0.981	0.997	$1.004^{1.048,\ 1.090}_{0.978,\ 0.950}$	$1.020^{1.063,\ 1.109}_{0.989,\ 0.957}$
$\log(10^{10}A_s)$	3.176	3.169	$3.138^{3.198,\ 3.250}_{3.065,\ 2.979}$	$3.124^{3.182,\ 3.236}_{3.038,\ 2.944}$
R	0.003	0.017	$0.000^{0.107,\ 0.237}_{0.000,\ 0.000}$	$0.000^{0.131,\ 0.274}_{0.000,\ 0.000}$
Run 3 (varying τ, n_s^{low}, n_s^{high}, and A_s)				
τ	0.0894	0.0924	$0.0906^{0.1086,\ 0.1283}_{0.0725,\ 0.0573}$	$0.0929^{0.1069,\ 0.1240}_{0.0761,\ 0.0624}$
n_s^{low}	1.025	1.039	$1.031^{1.096,\ 1.162}_{0.960,\ 0.893}$	$1.025^{1.105,\ 1.178}_{0.966,\ 0.898}$
n_s^{high}	1.022	1.032	$1.021^{1.049,\ 1.073}_{0.999,\ 0.975}$	$1.034^{1.055,\ 1.080}_{1.005,\ 0.980}$
$\log(10^{10}A_s)$	3.105	3.093	$3.109^{3.155,\ 3.197}_{3.061,\ 3.014}$	$3.106^{3.146,\ 3.189}_{3.052,\ 3.006}$
Run 4 (varying τ, A_s, and R)				
τ	0.0726	0.0830	$0.0722^{0.0868,\ 0.1010}_{0.0588,\ 0.0460}$	$0.0825^{0.0943,\ 0.1067}_{0.0687,\ 0.0554}$
$\log(10^{10}A_s)$	3.257	3.230	$3.260^{3.285,\ 3.312}_{3.228,\ 3.200}$	$3.226^{3.251,\ 3.275}_{3.199,\ 3.171}$
R	4×10^{-5}	2×10^{-4}	$0.000^{0.011,\ 0.029}_{0.000,\ 0.000}$	$0.000^{0.018,\ 0.043}_{0.000,\ 0.000}$
Run 5 (varying n_s^{low}, n_s^{high}, and A_s)				
n_s^{low}	1.009	1.022	$1.013^{1.076,\ 1.141}_{0.943,\ 0.882}$	$1.019^{1.080,\ 1.153}_{0.952,\ 0.894}$
n_s^{high}	1.017	1.024	$1.018^{1.038,\ 1.060}_{0.994,\ 0.973}$	$1.022^{1.046,\ 1.068}_{1.000,\ 0.979}$
$\log(10^{10}A_s)$	3.101	3.089	$3.100^{3.148,\ 3.189}_{3.058,\ 3.012}$	$3.088^{3.137,\ 3.179}_{3.045,\ 3.000}$
Run 6 (varying A_s and R)				
$\log(10^{10}A_s)$	3.271	3.225	$3.274^{3.280,\ 3.288}_{3.262,\ 3.254}$	$3.223^{3.232,\ 3.240}_{3.215,\ 3.207}$
R	8×10^{-6}	2×10^{-5}	$0.000^{0.012,\ 0.031}_{0.000,\ 0.000}$	$0.000^{0.019,\ 0.044}_{0.000,\ 0.000}$

Table 9.3: Unmarginalised ML points and one-dimensional peaks and MCI limits resulting from a likelihood analysis using the modified version of CosmoMC to include an l-dependent n_s. Both 5-year and 7-year $WMAP$ data sets were used, and all runs were over the range $l = 2 - 220$. MCI limits are given with both 68% and 95% confidence.

Parameter	ML value		One-dimensional peaks and MCI limits $X^{\,68\%\uparrow,\ 95\%\uparrow}_{\,68\%\downarrow,\ 95\%\downarrow}$	
	5-year	7-year	5-year	7-year
Run 1 (varying n_s^{low}, n_s^{high}, A_s, and R)				
n_s^{low}	0.963	0.965	$0.945^{\,1.002,\ 1.052}_{\,0.896,\ 0.843}$	$0.960^{\,1.004,\ 1.057}_{\,0.896,\ 0.847}$
n_s^{high}	0.959	0.963	$0.962^{\,0.971,\ 0.978}_{\,0.954,\ 0.946}$	$0.965^{\,0.974,\ 0.981}_{\,0.958,\ 0.950}$
$\log(10^{10}A_s)$	3.213	3.209	$3.206^{\,3.223,\ 3.249}_{\,3.183,\ 3.161}$	$3.197^{\,3.221,\ 3.243}_{\,3.178,\ 3.159}$
R	6×10^{-5}	4×10^{-4}	$0.000^{\,0.040,\ 0.092}_{\,0.000,\ 0.000}$	$0.000^{\,0.042,\ 0.090}_{\,0.000,\ 0.000}$
Run 2 (varying τ, n_s^{low}, n_s^{high}, A_s, and R)				
τ	0.0795	0.0821	$0.0803^{\,0.0966,\ 0.1123}_{\,0.0669,\ 0.0520}$	$0.0808^{\,0.0953,\ 0.1096}_{\,0.0684,\ 0.0560}$
n_s^{low}	0.960	0.973	$0.950^{\,1.001,\ 1.051}_{\,0.896,\ 0.845}$	$0.953^{\,1.007,\ 1.057}_{\,0.899,\ 0.847}$
n_s^{high}	0.960	0.964	$0.962^{\,0.970,\ 0.979}_{\,0.954,\ 0.946}$	$0.967^{\,0.974,\ 0.982}_{\,0.958,\ 0.950}$
$\log(10^{10}A_s)$	3.209	3.206	$3.206^{\,3.241,\ 3.273}_{\,3.174,\ 3.140}$	$3.201^{\,3.234,\ 3.264}_{\,3.172,\ 3.140}$
R	4×10^{-4}	0.003	$0.000^{\,0.039,\ 0.087}_{\,0.000,\ 0.000}$	$0.000^{\,0.043,\ 0.095}_{\,0.000,\ 0.000}$
Run 3 (varying τ, n_s^{low}, n_s^{high}, and A_s)				
τ	0.0788	0.0778	$0.0751^{\,0.0922,\ 0.1088}_{\,0.0626,\ 0.0497}$	$0.0798^{\,0.0918,\ 0.1045}_{\,0.0660,\ 0.0534}$
n_s^{low}	0.913	0.924	$0.920^{\,0.970,\ 1.020}_{\,0.865,\ 0.813}$	$0.911^{\,0.971,\ 1.017}_{\,0.866,\ 0.816}$
n_s^{high}	0.966	0.970	$0.966^{\,0.976,\ 0.983}_{\,0.958,\ 0.950}$	$0.968^{\,0.977,\ 0.985}_{\,0.961,\ 0.953}$
$\log(10^{10}A_s)$	3.190	3.182	$3.185^{\,3.220,\ 3.254}_{\,3.154,\ 3.124}$	$3.185^{\,3.217,\ 3.246}_{\,3.157,\ 3.127}$
Run 4 (varying τ, A_s, and R)				
τ	0.0785	0.0815	$0.0794^{\,0.0912,\ 0.1058}_{\,0.0629,\ 0.0492}$	$0.0786^{\,0.0922,\ 0.1045}_{\,0.0662,\ 0.0546}$
$\log(10^{10}A_s)$	3.223	3.234	$3.217^{\,3.248,\ 3.277}_{\,3.191,\ 3.164}$	$3.226^{\,3.255,\ 3.279}_{\,3.204,\ 3.179}$
R	1×10^{-4}	4×10^{-4}	$0.000^{\,0.016,\ 0.039}_{\,0.000,\ 0.000}$	$0.000^{\,0.016,\ 0.040}_{\,0.000,\ 0.000}$
Run 5 (varying n_s^{low}, n_s^{high}, and A_s)				
n_s^{low}	0.925	0.923	$0.911^{\,0.968,\ 1.013}_{\,0.868,\ 0.822}$	$0.938^{\,0.978,\ 1.027}_{\,0.873,\ 0.818}$
n_s^{high}	0.968	0.970	$0.965^{\,0.974,\ 0.982}_{\,0.958,\ 0.952}$	$0.969^{\,0.978,\ 0.985}_{\,0.962,\ 0.954}$
$\log(10^{10}A_s)$	3.189	3.189	$3.190^{\,3.216,\ 3.232}_{\,3.173,\ 3.151}$	$3.191^{\,3.210,\ 3.230}_{\,3.166,\ 3.147}$
Run 6 (varying A_s and R)				
$\log(10^{10}A_s)$	3.298	3.232	$3.297^{\,3.304,\ 3.310}_{\,3.292,\ 3.286}$	$3.232^{\,3.237,\ 3.242}_{\,3.227,\ 3.222}$
R	2×10^{-5}	2×10^{-5}	$0.000^{\,0.008,\ 0.021}_{\,0.000,\ 0.000}$	$0.000^{\,0.015,\ 0.038}_{\,0.000,\ 0.000}$

Table 9.4: Unmarginalised ML points and one-dimensional peaks and MCI limits resulting from a likelihood analysis using the modified version of CosmoMC to include an l-dependent n_s. Both 5-year and 7-year $WMAP$ data sets were used, and all runs were over the range $l = 2 - l_{max}$. MCI limits are given with both 68% and 95% confidence.

Parameters varied	$\log \mathcal{L}$ (5-year)	$\log \mathcal{L}$ (7-year)
n_s, A_s, and R (original analysis)	284.938	134.985
n_s and A_s (original analysis)	284.244	134.622
n_s^{low}, n_s^{high}, A_s, and R	284.972	135.044
n_s^{low}, n_s^{high}, and A_s	284.243	134.622
A_s and R	275.896	131.768
τ, n_s, A_s, and R (original analysis)	284.981	135.115
τ, n_s, and A_s (original analysis)	284.386	134.832
τ, n_s^{low}, n_s^{high}, A_s, and R	285.027	135.166
τ, n_s^{low}, n_s^{high}, and A_s	284.392	134.816
τ, A_s, and R	276.034	131.739

Table 9.5: Log(likelihood) values of the ML points of various runs from both the first and second analyses. The l-range is always 2-220.

model found fits only marginally better with data than the constant-n_s model for the same l range.

In order to shed further light on why this may be so, we ask a different, but related question: Why does the ML and one-dimensional peak values for n_s^{low} change when we go from $l = 2 - 220$ to $2 - l_{max}$ (for instance, compare run 1 in tables 9.3 and 9.4)? One might perhaps think that since n_s^{low} was only operative for $l < 100$, the inclusion of more l's beyond 220 should have no effect on this quantity. We should then once again recall the parameter degeneracies we saw in chapter 6, fig. 9.2, and for this analysis, fig. 9.3. For a given l range, the data puts certain constraints on the relevant parameters, and gives certain ML values. Increasing this range, we may find that there is a value for, e.g., A_s that fits the small-scale anisotropies better. In that case, we must change other parameters, like n_s^{low} or τ, in order to make the large-scale spectrum fit again. This way, even the parameters that have nothing to do directly with some data set may nevertheless be influenced indirectly by that data set through the more accurate determination of other parameters. This also explains why we do not get the same values for n_s^{low} and n_s^{high} as the values for n_s found in the $l = 2 - 100$ and $l = 101 - 220$ runs, respectively: The union of both l ranges introduce constraints that are not present in the ranges by themselves - constraints which indirectly also change the spectral indices from what they otherwise would have been.

9.6 Chapter summary

In this chapter, we have presented results from two separate analyses: The repetition of the analysis in Zhao et al. (2009b), albeit with certain improvements, and the testing of the claim made in that paper that the low R values that follows from a likelihood analysis for $l \leq 220$ are due to the assumed constancy of n_s. We have seen that our first analysis largely verifies the results of the original analysis, except the inconsistencies between the one-dimensional distributions for n_s obtained using $l = 2 - 100$ and $l = 101 - 220$: We found that these distributions overlap within a 1σ area. As for our second analysis, we found that it to a large extent disproved the abovementioned claim, using a steplike n_s to test it. More specifically, this steplike spectral index did not yield an R value comparable to the value found in the $l = 2 - 100$ analyses, and fixing either R or n_s did not give the expected results.

We conclude that the assumed constancy of n_s, whether this assumption is erroneous or not, can not be the source of the low tensor contribution that arises when all the data are taken into account. However, even though the tensor contribution is low, it does not mean it is necessarily zero: We saw 'hints' on a $> 1\sigma$ level for one run using the $l = 2 - 100$ data that there is a nonzero contribution from tensors. Our second analysis with the full l range taken into account gave a negligible 68% upper limit on R, but this analysis was limited in that it did not take into account other parameters, in addition to implementing a nonstandard l-dependent spectral index. From our more complete analysis in sec. 9.4, we do find support for a nonzero gravitational wave contribution, even though the most likely value of r is zero. All of the above, of course, must again be qualified by the discussion in sec. 9.3.1: The hints that we have found are based on the assumed correctness of the fixed parameters.

Finally, in order to make contact with the previous chapter: It is clear that in none of this we find support for eq. (8.29) or even a dominant tensor contribution over the scalar one. We may thus also rule out Grishchuk's claim from that chapter observationally.

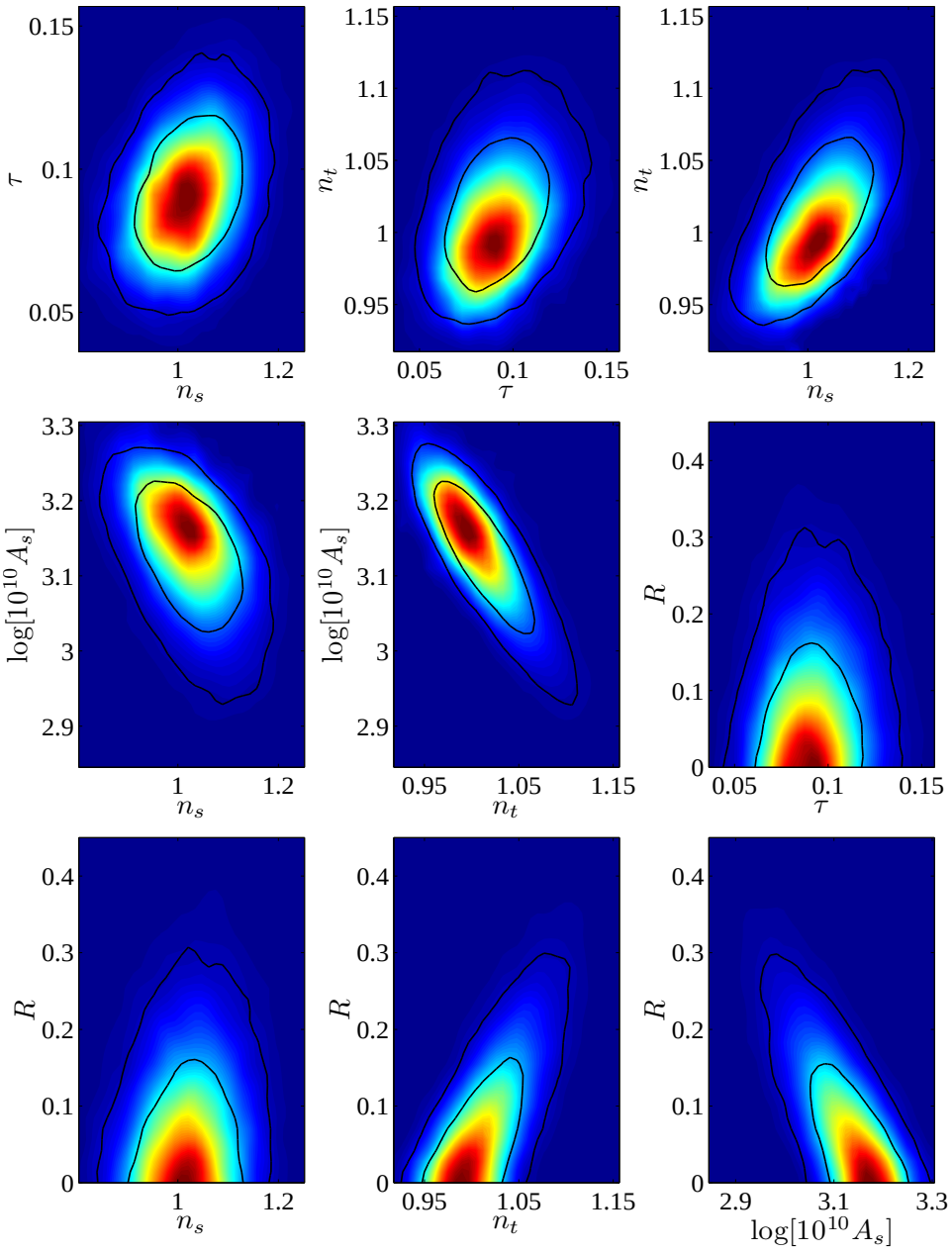

Figure 9.3: Two-dimensional marginalised distributions for the $l = 2 - 220$ $WMAP$ data, using a step-like spectral index. The solid lines are the 68% and 95% contours, and the colors indicate the mean likelihood value at that point.

Chapter 10

Conclusion

In this text, we have mainly studied the gravitational waves that were present in the early Universe, and the effect that these are expected to have on the Cosmic Microwave Background. As a consequence, we have also been studying the scalar (density) perturbations, since these also have an effect on the CMB, and keeping the two contributions apart can be tricky. We have examined the controversy that began with L. P. Grishchuk's paper Grishchuk (1994), and the claims of the authors of Zhao et al. (2009a) and Zhao et al. (2009b).

In chapter 2 and Part I, we reviewed the Standard Model of cosmology, and found that the standard theory predicts that the primordial power spectrum of gravitational waves should be negligible compared to the power spectrum of density perturbations. We also saw that the primordial perturbations have an effect on the CMB, and examined the various intensity and polarisation spectras that would arise from models with various gravitational wave contributions, and the degeneracy of these contributions with other primordial and cosmological parameters. We saw that gravitational waves give rise to a nonzero BB spectrum, but that the large noise values for this mode so far prevents detection of such a spectrum.

In Part II, we first summarised the Grishchuk controversy, which in a large part revolved around his claim that the contribution of gravitational waves should be non-negligible in the primordial epoch, possibly even greater than the scalar contribution.

We then moved on to the details of the controversy and tried to clear up some of it. We found that it seemed like Grishchuk had been wrong in his claim, and that he at least once changed his argumentation so that it first was the reheating transition that, according to him, was incorrectly treated,

and then the initial conditions of the early Universe. We found that both lines of argumentation failed, and that there is little theoretical reason to abandon the predictions of the Standard Model.

After this, we reviewed the two papers by Zhao et. al. and performed the likelihood analysis in Zhao et al. (2009b) with several improvements over the analysis in that paper. We also tested their claim that the negligible contribution of gravitational waves reported by the $WMAP$ team was due to the assumed constancy of n_s over the whole l range. This was done by implementing a 'step-like' spectral index n_s, which could assume two different values above and below $l = 100$. We found that our improved analysis did not change the original results much, except that the values of n_s for $l = 2-100$ and $l = 101-220$ were consistent with each other within a 68% confidence interval, in contrast to the original results. Further, we found that the assumed constancy of n_s could not have been solely responsible for the negligible contribution from gravitational waves.

With all this in mind, we again turn to the questions asked in chapter 1:

- It seems Grishchuk's claims about the incorrectness of the inflationary predictions are unfounded and based on small, but crucial, errors in the reasoning behind his mathematical formulas.
- The assumption of a constant spectral index can not be said to lead to a negligence of the signs of gravitational waves - if anything, the contrary is true. Our investigations have also shown that there is little reason to abandon this assumption in the first place.

Outlook

As we have seen, the small relative amount of gravitational waves in the early Universe has been the subject to some debate. Such debates are vital for any scientific enterprise; by sorting them out, the theory in question is either strengthened or weakened - its status seldom remains the same. They also serve as an important reminder that we must always revise our models and fix our critical eye upon them. It is critically important to be aware of the assumptions we make when testing our theories - are we really testing the theory, or are we framing the experiment in such a way that the theory is bound to be reinforced, no matter the observational outcome? We saw that the $WMAP$ analysis, for example, assumed the consistency relation of inflation to be true, even though one might argue that this relation should be verified by the data, not assumed beforehand - especially since the theory of inflation still remains a largely unverified theory as of yet. In order for the

theory of inflation to gain credit, we should let the data speak for themselves, not put words in their mouths - or at least be aware of when we are doing so.

The amount of tensor contributions to the CMB is just one of many predictions from inflation that may strengthen the theory if shown to be true. So far, we only have upper limits on this quantity. The *Planck* (Planck collaboration, 2005) and *QUIET* (Samtleben & QUIET collaboration, 2008) experiments currently in progress should provide much better constraints, and, as we saw earlier, possibly even a lower limit. The next generation of observational data will surely give us a much clearer picture of the Universe, and the role that gravitational waves played in its very first moments.

Appendix A

Conventions

- Natural units: $c = \hbar = 1$ where nothing else is noted.
- Metric signature: $\{-, +, +, +\}$.
- Fourier convention:

$$f(\vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} f(\vec{k})$$
$$f(\vec{k}) = \int d^3\vec{x} e^{-i\vec{k}\cdot\vec{x}} f(\vec{x})$$

  Note: Same symbol for the Fourier transformed function as for the original function.

- $\dot{x} = dx/d\eta$, $x' = \partial x/\partial t$, where η is conformal time, and t is cosmic time.
- Greek letters are four-indices, latin letters three-indices.
- A semicolon means that a covariant derivative is being taken, while a comma means that an ordinary partial derivative is being taken.

Appendix B

Miscellany

B.1 The Boltzmann equation

The evolution equations for the photon distribution perturbations are derived using the *Boltzmann equation*. This equation describes how the statistical distribution of a system of particles evolves with time in phase space. Given such a distribution $f(p, x, t)$, the Boltzmann equation, in a very simplified form, reads (Dodelson, 2003):

$$\frac{df}{dt} = C[f], \tag{B.1}$$

where C is some collision term depending on the distribution of the system and the particles of which it is made up.

B.2 Fourier transforms

Any mathematical function of a vector $f(\vec{x})$ satisfying certain criteria of smooth behaviour (Asmar, 2002) can be *Fourier transformed*:

$$f(x) = \frac{1}{(2\pi)^3} \int d^3k e^{i\vec{k}\cdot\vec{x}} g(\vec{k}) \tag{B.2}$$

where the integral goes over all k-space, and $g(\vec{k})$ is the *Fourier coefficient* of f. Linear relations between functions translate into linear relations between Fourier coefficients.

In addition, when f is a function of space, the wavenumber $k = |\vec{k}|$ can be interpreted as saying something about the scales which we are currently

studying: The Fourier coefficient with wavenumber k describes the amount of structure at the characteristic length scale k^{-1}. When talking about 'Fourier modes' at some wavenumber k, this is what we're referring to.

B.3 Spherical harmonic expansions

Let a sphere be described by a unit vector $\hat{p}$, and let h be any function on the sphere that is independent of the orientation of the tangent vectors at the point of evaluation. It can then be expanded using a *spherical harmonic* expansion (Bransden & Joachain, 2000):

$$h(\hat{p}) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\hat{p}), \tag{B.3}$$

where the Y_{lm} are the spherical harmonic functions[1], and the a_{lm}'s are the expansion coefficents, containing all information about the function h. Again, as with Fourier transforms, there is a characteristic scale associated with the expansion variable l - a higher l means smaller scales on the sphere.

Since the spherical harmonic functions obey certain orthonormality relations, we may turn eq. (B.3) around and express the a_{lm}'s in terms of the function h:

$$a_{lm} = \int d\Omega Y^*_{lm}(\hat{p}) h(\hat{p}). \tag{B.4}$$

B.4 Spin-weighted harmonics

The only restraint we put on a function on a sphere that can be expanded in spherical harmonics, was that it needed to be a scalar under *rotations of tangent vectors to the sphere* (which form a plane) at any point of evaluation. If, however, the function does not satisfy this, but rather that it transforms according to

$$f'_s(\hat{p}) = e^{-is\psi} f_s(\hat{p}) \tag{B.5}$$

under a rotation of the tangent vectors by an angle ψ, it is said to have spin s (Zaldarriaga & Seljak, 1997). Any function with nonzero spin cannot be expanded in spherical harmonics, so one must introduce the spin raising and

[1] By definition, the solutions to Laplace's equation in spherical coordinates.

lowering operators $\eth, \bar{\eth}$:

$$\begin{aligned}
\eth f_s(\theta,\phi) &= -\sin^s\left(\frac{\partial}{\partial\theta}+i\csc(\theta)\frac{\partial}{\partial\phi}\right)\sin^{-s}(\theta)f_s(\theta,\phi) \\
\bar{\eth} f_s(\theta,\phi) &= -\sin^{-s}(\theta)\left(\frac{\partial}{\partial\theta}-i\csc(\theta)\frac{\partial}{\partial\phi}\right)\sin^{s}(\theta)f_s(\theta,\phi), \qquad \text{(B.6)}
\end{aligned}$$

where we have switched to polar coordinates in our description of the sphere. Using these operators, once can apply the spin operators to the spherical harmonics, creating so-called *spin-weighted spherical harmonics* ${}_sY_{lm}$ and expanding the original function in these (Zaldarriaga & Seljak, 1997):

$$f_s(\hat{p}) = \sum_{lm} a_{s,lm}({}_sY_{lm}(\hat{p})). \qquad \text{(B.7)}$$

Alternatively, one may act with $\eth, \bar{\eth}$ on the spin-s-valued function, turning it into a spin-zero valued function, and then expanding the resulting function in ordinary spherical harmonics:

$$\begin{aligned}
(\bar{\eth})^s f_s(\hat{p}) &= \sum_{lm}\tilde{a}_{s,lm}Y_{lm}, \quad s>0 \\
(\eth)^s f_s(\hat{p}) &= \sum_{lm}\tilde{a}_{s,lm}Y_{lm}, \quad s<0, \qquad \text{(B.8)}
\end{aligned}$$

where the expansion coefficients $\tilde{a}_{s,lm}$ are related to the $a_{s,lm}$'s by

$$\tilde{a}_{s,lm} = \left(\frac{(l+s)!}{(l-s)!}\right)^{1/2} a_{s,lm}. \qquad \text{(B.9)}$$

Finally, the expansion coefficients can be calculated if the original function is known:

$$\begin{aligned}
a_{s,lm} &= \int d\Omega\, {}_sY^*_{lm}(\hat{p}) f_s(\hat{p}) \\
&= \left(\frac{(l+2)!}{(l-2)!}\right)^{-1/2}\int d\Omega Y^*_{lm}(\hat{p})(\bar{\eth})^s f_s(\hat{p}), \quad s>0 \\
&= \left(\frac{(l+2)!}{(l-2)!}\right)^{-1/2}\int d\Omega Y^*_{lm}(\hat{p})(\eth)^s f_s(\hat{p}), \quad s<0 \qquad \text{(B.10)}
\end{aligned}$$

B.5 The polarisation of electromagnetic waves

Consider, very generally, an electromagnetic wave travelling in the z-direction (we can always choose our coordinate system in this way). Its electric field vector can then be written as (Kosowsky, 1996):

$$E_i = a_i(t)\cos(\omega_0 t - \theta_i(t)), \qquad \text{(B.11)}$$

where $i = x, y$, ω_0 is the frequency of the wave, a_i are the amplitudes of the wave, and θ_i are the wave's phases. We are assuming that the wave is nearly monochromatic: its amplitudes and phases are slowly varying compared to the inverse frequency.

The wave is said to be *polarised* if some correlation exists between the electric field in the x and y direction. By taking the time averages over the field amplitudes a_i, we can define the *Stokes parameters*:

$$\begin{aligned} I &\equiv \langle a_x^2 \rangle + \langle a_y^2 \rangle \\ Q &\equiv \langle a_x^2 \rangle - \langle a_y^2 \rangle \\ U &\equiv \langle 2a_x a_y \cos(\theta_x - \theta_y) \rangle \\ V &\equiv \langle 2a_x a_y \sin(\theta_x - \theta_y) \rangle \end{aligned} \tag{B.12}$$

The Stokes parameter I is just the intensity of the electromagnetic wave, while the other parameters describe the polarisation of the wave: Q and U describe linear polarisation, and V describes circular polarisation. If $Q = U = V = 0$, the wave is unpolarised.

The intensity I and circular polarisation V are invariant under rotations of the coordinate system, but Q and U are not: If the x and y axes are rotated by an angle ψ, the parameters Q and U can be shown to transform as (Zaldarriaga & Seljak, 1997)

$$\begin{aligned} Q' &= Q\cos(2\psi) + U\sin(2\psi) \\ U' &= -Q\sin(2\psi) + U\cos(2\psi) \\ (Q \pm iU)' &= e^{\pm 2i\psi}(Q \pm iU). \end{aligned} \tag{B.13}$$

An alternative definition of polarisation based on quantum mechanical concepts is as follows (Kosowsky, 1996): Let $|\epsilon_x\rangle$ and $|\epsilon_y\rangle$ be a pair of orthonormal basis vectors in the photon polarisation state space, and take them to represent the x and y direction, respectively. Any photon polarisation state can then be written as

$$|\epsilon\rangle = a_x e^{i\theta_x}|\epsilon_x\rangle + a_y e^{i\theta_y}|\epsilon_y\rangle, \tag{B.14}$$

where $\theta_i, i = x, y$, are arbitrary phases, and a_i are the probability amplitudes corresponding to each state. The Stokes parameter operators are then given by

$$\begin{aligned} \hat{I} &= |\epsilon_x\rangle\langle\epsilon_x| + |\epsilon_y\rangle\langle\epsilon_y| \\ \hat{Q} &= |\epsilon_x\rangle\langle\epsilon_x| - |\epsilon_y\rangle\langle\epsilon_y| \\ \hat{U} &= |\epsilon_x\rangle\langle\epsilon_y| + |\epsilon_y\rangle\langle\epsilon_x| \\ \hat{V} &= i|\epsilon_y\rangle\langle\epsilon_x| - i|\epsilon_x\rangle\langle\epsilon_y| \end{aligned} \tag{B.15}$$

and the polarisation density matrix ρ is given by

$$\rho = \frac{1}{2}\begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix}, \tag{B.16}$$

where I, Q, U and V are the expectation values of the Stokes parameter operators.

B.6 Quantum field theory

This section will review, in as much detail as is needed, the quantum theory of fields. Our approach will be to start with the classical field theory and then move on to the quantum version of it. All of the information in this section is from Lyth & Liddle (2009).

B.6.1 Field theory

We start by recalling that in classical mechanics, all laws of physics can be derived from the *action*

$$S = \int L dt, \tag{B.17}$$

where t is some time coordinate and L is the *Lagrangian*, a function containing all relevant information about the physical system one is studying. It is generally a function of some degrees of freedom q_n and their time derivatives q_n', which may or may not be the actual three-dimensional coordinates. In addition, the Lagrangian depends on the time t. The evolution equations for these degrees of freedom are then found from the *action principle*, which says that a general variation of the action should be zero: $\delta S = 0$. This leads to the *Euler-Lagrange equations*

$$\frac{\partial L}{\partial q_n} - \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial q_n'}\right) = 0 \tag{B.18}$$

which are valid for any physical system appropriately described by some Lagrangian.

We then move to field theory, and introduce the concept of fields. According to field theory, every spacetime point can be associated with one or more fields (we shall work with only one for simplicity), which then become the degrees of freedom in the Lagrangian (since there is a field value for every spacetime point, the number of degrees of freedom becomes infinite).

Because the action must be Lorentz invariant to satisfy the principle of relativity, the Lagrangian can be written as $L = \int \mathcal{L} d^3x$, where $\mathcal{L}$ is called the *Lagrangian density*. The Lagrangian is then not a function of the spatial coordinates.

The energy-momentum tensor for a field may be found from its Lagrangian density using the following relation:

$$T_{\mu\nu} = -2\frac{\partial \mathcal{L}}{\partial g^{\mu\nu}} + g_{\mu\nu}\mathcal{L}. \tag{B.19}$$

B.6.2 Quantisation

We now incorporate the concepts of states and operators from quantum mechanics. We work in the Heisenberg picture, in which the states are time independent, while the operators evolve in time according to the Heisenberg equation of motion. We promote the degrees of freedom to operator status, so that the Lagrangian (and/or Lagrangian density) becomes an operator function, $L = L(t, \hat{q}_n, \hat{q}'_n)$. By promoting our fields to operators, we have essentially arrived at quantum field theory (though the details of this are more gritty than we let on here). It turns out that the fields can be associated with particles, and that by postulating different transformation properties for the fields under Lorentz transformations, one may describe different types of particles, with different spin values.

It is important to note that all of the above is done in the context of a flat spacetime - for curved spacetimes, some of the details change, but the flat spacetime approximation will be a fine approximation for everything we do in this text.

B.7 Markov Chains

Let $\vec{x}$ be a state vector for some system, and suppose the state of the system changes over time in discrete iterations. If the exact state of the system at any future time cannot be known exactly, but the probability for a given state to occur is known completely if the present state is known, the process is called a *Markov process* (or *chain*) (Anton & Rorres, 2005).

In order to know the probability distribution function for $\vec{x}$ at iteration $t+1$, $p^{t+1}(\vec{x})$, we must have knowledge of the *transition matrix* T. Once this is known, the PDF of $\vec{x}'$ at iteration $t+1$ becomes

$$p^{t+1}(\vec{x}') = \int d^N\vec{x} T(\vec{x}'; \vec{x}) p^t(\vec{x}), \tag{B.20}$$

where N is the dimensionality of the state vector. This can be applied iteratively, so that all one needs to know in addition to the transition matrix is some initial probability distribution $p^0(\vec{x})$ (McKay, 2003).

An *equilibrium distribution* π for a Markov Chain is one that satisfies

$$\pi(\vec{x}') = \int d^N \vec{x} T(\vec{x}'; \vec{x}) \pi(\vec{x}). \tag{B.21}$$

Appendix C

Modifications to CosmoMC

In this appendix, we provide the changes to CosmoMC that were carried out for the data analysis in chapter 9. Since the data analysis had two parts to it, so will this appendix - one for the code that implements the tensor-to-scalar quadrupole ratio R, and one for the code that implements the step-like spectral index n_s.

C.1 Implementing R

In order to use the tensor-to-scalar quadrupole ratio R instead of the primordial power spectrum ratio r, we need to make the following changes to the CosmoMC code:

In the `camb` folder, we modify the the following files in the following ways: In the file `modules.f90`, we define the parameter `Zrat` in the `CAMBParams` type:

```
type CAMBParams
    ...
    real(dl) Zrat
    ...
end type CAMBParams
```

This parameter is R. We also define the following subroutine in the same file:

```
subroutine ensure_ratio
use precision
use ModelParams
    integer in
    real(dl) fac

    do in =1, CP%InitPower%nn
        if (CP%WantTensors) then
            fac = Cl_scalar(2, in, C_Temp)/Cl_tensor(2,in, CT_Temp)&
            *CP%Zrat
            Cl_tensor(lmin:CP%Max_l_tensor, in, CT_Temp:CT_Cross) = &
            Cl_tensor(lmin:CP%Max_l_tensor, in, CT_Temp:CT_Cross)*fac
        end if
    end do

end subroutine ensure_ratio
```

The `CP%Initpower%nn` parameter is always 1 in our analyses. This subroutine normalizes the spectra so that the II tensor-to-scalar quadrupole ratio is equal to R.

Then, in the file `camb.f90`, we insert calls to the above subroutine:

```
subroutine CAMB_TransfersToPowers(CData)
    ...
    if (CData%Params%WantCls) then
        call ClTransferToCls(...)
        call ensure_ratio !This is our modification
        if (CP%DoLensing) call lens_Cls
        ...
    end if
    ...
end subroutine CAMB_TransfersToPowers
...
subroutine CAMB_GetResults(Params, error)
    ...
    if (.not. CP%OnlyTransfers) then
        if CP%WantCls call ensure_ratio !our modification
        !if (CP%WantCls .and. CP%OutputNormalization == outCOBE) call &
        COBEnormalize !We comment this bit out, since we don't use this
                      !normalization
        if (CP%DoLensing) then
            call lens_Cls
```

```
        end if
    end if
end subroutine CAMB_GetResults
```

We must insert the call twice because which of the two subroutines is used to get the power spectrum varies. Note that we in both instances insert the normalization *before* lensing is carried out - which makes sense since R is essentially a primordial parameter, while lensing happens after primordial times, so that fixing R after lensing would make less physical sense.

These are all the changes that are necessary for the `camb` folder, and thus the part of the program that actually computes the power spectrum for a given model. We must also make changes in the `source` folder, which contains the MCMC sampler and the part that relates the CosmoMC parameters to the $CAMB$ parameters. So, we first change the file `CMB_Cls_simple.f90`:

```
subroutine CMBToCAMB(CMB, P)
    ...
    type(CAMBParams) P
    ...
    P%Zrat = CMB%norm(norm_amp_ratio)
    ...
end subroutine CMBToCAMB
```

Then, we change `params_CMB.f90`:

```
subroutine SetCAMBInitPower(P, CMB, in)
    ...
    type(CAMBParams) P
    ...

    !P%InitPower%rat(in) = CMB%norm(norm_amp_ratio) !We comment this out
    P%InitPower%rat(in) = 1.d0 !... and insert this instead
    P%Zrat = CMB%norm(norm_amp_ratio)
    ...
    !P%InitPower%ant(in) = CMB%InitPower(2) !Again, we comment this out
    P%InitPower%ant(in) = P%InitPower%an(in) -1.d0 !And use this instead
    ...
end subroutine SetCAMBInitPower(P, CMB, in)
```

Again, CosmoMC sometimes uses `CMBToCAMB`, other times `SetCAMBInitPower` to relate the CosmoMC parameters to the $CAMB$ parameters, so we must

change both identically. In addition, we here have set r equal to 1 (which doesn't matter, since we will normalize the power spectrum later), letting R take its place. The last modification sets n_t equal to $n_s - 1$, as mentioned above.

These are the modifications done in order to use R as a parameter in the CosmoMC software instead of r.

C.2 Implementing a step-like n_s

We here provide the code changes used to implement a step-like spectral index in CosmoMC, with a step at $l \approx 100$. These changes come in addition to the changes already described in the previous section.

We again begin with the `camb` folder, in which we change the file `power_tilt.f90`:

```
Type InitialPowerParams
    ...
    !real(dl) an(nnmax) !Comment this out (scalar spectral indices)
    real(dl) an(2, nnmax)
    ...
    !real(dl) ant(nnmax) !The same (tensor spectral indices)
    real(dl) ant(2, nnmax)
    ...
    real(dl) l_step !Contains the l value at which the step function steps
end Type InitialPowerParams
...
!function ScalarPower(k, in) !Comment this out
function ScalarPower(k, in, omegam, H0)
    ...
    !real(dl) ScalarPower, k, lnrat !Comment this out
    real(dl) ScalarPower, k, lnrat, omegam, H0, lightspeed, d_a, k_step
    ...
    lightspeed = 2992792.458d0
    d_a = 2.d0*lightspeed/(H0*(omegam)**(0.4d0))
    k_step = dble(P%lstep)/d_a !P is an InitialPowerParams variable
    ...
    !ScalarPower=P%ScalarPowerAmp(in)*exp( (P%an(in)-1)*lnrat + &
      P%n_run(in)/2*lnrat**2) !Comment this out
    if (k < k_step) then
        ScalarPower = P%ScalarPowerAmp(in)*exp((P%an(1, in) -1)*lnrat + &
          P%n_run(in)/2*lnrat**2)
```

```
    else
        ScalarPower = P%ScalarPowerAmp(in)*exp((P%an(2, in) -1)*lnrat + &
          P%n_run(in)/2*lnrat**2)
    end if
end function ScalarPower
```

There is a similar function, called `TensorPower`, which was modified in exactly the same way. These are the main changes to the `camb` folder. In addition, all references to `an, ant, ScalarPower` or `TensorPower` were modified so that they contained the correct number of arguments or indices.

In the `source` folder, the abovementioned changes to the modified variables and functions were also carried out. In addition, the following changes were made to the `params_CMB.f90` file:

```
subroutine SetCAMBInitPower(P, CMB, in)
    ...
    P%InitPower%l_step = 100
    ...
    !P%InitPower%an(in) = CMB%InitPower(1) !Comment this out
    P%InitPower%an(1, in) = CMB%InitPower(1)
    P%InitPower%an(2, in) = CMB%InitPower(2)
    ...
    !P%InitPower%ant(in) = CMB%InitPower(2) !Comment this out
    P%InitPower%an(1, in) = P%InitPower%an(1, in) - 1.d0
    P%InitPower%an(2, in) = P%InitPower%an(2, in) - 1.d0
ens subroutine SetCAMBInitPower
```

This lets n_s^{high} take the role of n_t, while n_s^{low} takes the n_s role. We then set n_t (high and low) to be $n_s - 1$ (high and low).

Index

Bibliography

Adelman-McCarthy, J. K. et al.: 2006, "The Fourth Data Release of the Sloan Digital Sky Survey", *Astrophys. J. Suppl. S.* **162**, 38–48, arXiv:astro-ph/0507711

Anton, H., and C. Rorres: 2005, *Elementary Linear Algebra: Applications Version*, John Wiley & Sons, Inc., Hoboken, New Jersey, ninth edition

Asmar, N. H.: 2002, *Applied Complex Analysis with Partial Differential Equations*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, first edition

Balland, C. et al.: 2009, "The ESO/VLT 3rd year Type Ia supernova data set from the supernova legacy survey", *Astron. Astrophys.* **507**, 85–103, arXiv:0909.3316

Bardeen, J. M.: 1980, "Gauge-invariant cosmological perturbations", *Phys. Rev. D* **22**, 1882–1905

Baumann, D. et al.: 2009, "Probing Inflation with CMB Polarization", in *American Institute of Physics Conference Series*, American Institute o63/1.3160885

Becker, R. H. et al.: 2001, "Evidence for Reionization at z∼6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar", *Astron. J.* **122**, 2850–2857, arXiv:astro-ph/0108097

Bernardo, J. M.: 2003, "Bayesian statistics",, in R. Viertl (Ed.), *Encyclopedia of Life Support Systems (EOLSS). Probability and Statistics*, Oxford, UK

Bond, J. R., and G. Efstathiou: 1984, "Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter", *Astrophys. J. Lett.* **285**, L45–L48

Bransden, B. H., and C. J. Joachain: 2000, *Quantum Mechanics*, Pearson Education, Ltd., Harlow, Essex, second edition

Caldwell, R. R.: 1996, "On the evolution of scalar metric perturbations in an inflationary cosmology", *Classical and Quantum Gravity* **13**, 2437–2447, arXiv:gr-qc/9509027

Christensen, N., R. Meyer, L. Knox, and B. Luey: 2001, "Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements", *Classical and Quantum Gravity* **18**, 2677–2688, arXiv:astro-ph/0103134

Colombo, L. P. L., E. Pierpaoli, and J. R. Pritchard: 2009, "Cosmological parameters after WMAP5: forecasts for Planck and future galaxy surveys", *Mon. Not. R. Astron. Soc.* **398**, 1621–1637, arXiv:0811.2622

Crittenden, R., J. R. Bond, R. L. Davis, G. Efstathiou, and P. J. Steinhardt: 1993, "Imprint of gravitational waves on the cosmic microwave background", *Physical Review Letters* **71**, 324–327, arXiv:astro-ph/9303014

Deruelle, N., and V. F. Mukhanov: 1995, "Matching conditions for cosmological perturbations", *Phys. Rev. D* **52**, 5549–5555, arXiv:gr-qc/9503050

Dodelson, S.: 2003, *Modern Cosmology*, Academic Press, California, first edition

Dodelson, S., and J. M. Jubas: 1995, "Reionization and its imprint of the cosmic microwave background", *Astrophys. J.* **439**, 503–516, arXiv:astro-ph/9308019

Elgarøy, Ø., M. Gramann, and O. Lahav: 2002, "Features in the primordial power spectrum: constraints from the cosmic microwave background and the limitation of the 2dF and SDSS redshift surveys to detect them", *Mon. Not. R. Astron. Soc.* **333**, 93–99, arXiv:astro-ph/0111208

Fan, X. et al.: 2002, "New z $>$ 6 Quasars Discovered in the Sloan Digital Sky Survey", in *Bulletin of the American Astronomical Society*, Vol. 34 of *Bulletin of the American Astronomical Society*, 1288–+

Gotz, M.: 1998, "Validity of the 'conservation law' in the evolution of cosmological perturbations", *Mon. Not. R. Astron. Soc.* **295**, 873–876, arXiv:astro-ph/9704271

Grishchuk, L. P.: 1974, "Amplification of gravitational waves in an isotropic universe", *Soviet Journal of Experimental and Theoretical Physics* **40**, 409–+

Grishchuk, L. P.: 1993a, "Cosmological rotation of quantum-mechanical origin and anisotropy of the microwave background", *Phys. Rev. D* **48**, 5581–5593, arXiv:gr-qc/9310011

Grishchuk, L. P.: 1993b, "Quantum effects in cosmology", *Classical and Quantum Gravity* **10**, 2449–2477, arXiv:gr-qc/9302036

Grishchuk, L. P.: 1994, "Density perturbations of quantum-mechanical origin and anisotropy of the microwave background", *Phys. Rev. D* **50**, 7154–7172, arXiv:gr-qc/9405059

Grishchuk, L. P.: 1995, "Comment on the paper of Leonard Parker and Yang Zhang "Cosmological perturbations of a relativistic condensate"",, arXiv:gr-qc/9506010

Grishchuk, L. P.: 1996, "Statistics of the microwave background anisotropies caused by cosmological perturbations of quantum-mechanical origin.", in N. Sanchez & A. Zichichi (Ed.), *String Gravity and Physics at the Planck Energy Scale*, 369–408

Grishchuk, L. P.: 1998, "Comment on the "Influence of Cosmological Transitions on the Evolution of Density Perturbations"",, arXiv:gr-qc/9801011

Grishchuk, L. P.: 2005, "REVIEWS OF TOPICAL PROBLEMS: Relic gravitational waves and cosmology", *Physics Uspekhi* **48**, 1235–1247, arXiv:gr-qc/0504018

Grishchuk, L. P., and I. B. Zeldovich: 1978, "Long-wavelength perturbations of a Friedmann universe, and anisotropy of the microwave background radiation", *Soviet Astronomy* **22**, 125–129

Grøn, Ø., and S. Hervik: 2007, *Einstein's General Theory of Relativity*, Springer, New York, first edition

Hamann, J., S. Hannestad, G. G. Raffelt, and Y. Y. Y. Wong: 2007, "Observational bounds on the cosmic radiation density", *Journal of Cosmology and Astro-Particle Physics* **8**, 21–+, arXiv:0705.0440

Hicken, M. et al.: 2009, "Improved Dark Energy Constraints from ˜100 New CfA Supernova Type Ia Light Curves", *Astrophys. J.* **700**, 1097–1140, arXiv:0901.4804

Komatsu, E. et al.: 2009, "Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation", *Astrophys. J. Suppl. S.* **180**, 330–376, arXiv:0803.0547

Komatsu, E. et al.: 2010, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation",, arXiv:1001.4538

Kosowsky, A.: 1996, "Cosmic microwave background polarization.", *Annals of Physics* **246**, 49–85, arXiv:astro-ph/9501045

Lewis, A.: 2010, *CosmoMC Readme*, `http://cosmologist.info/cosmomc/readme.html`

Lewis, A., and S. Bridle: 2002, "Cosmological parameters from CMB and other data: a Monte-Carlo approach", *Phys. Rev.* **D66**, 103511, astro-ph/0205436

Lewis, A., and A. Challinor: 2006, "Weak gravitational lensing of the CMB", *Phys. Rep.* **429**, 1–65, arXiv:astro-ph/0601594

Lewis, A., A. Challinor, and A. Lasenby: 2000, "Efficient Computation of CMB anisotropies in closed FRW models", *Astrophys. J.* **538**, 473–476

Liddle, A. R.: 1999, "An Introduction to Cosmological Inflation", in R. L. Walsworth & D. F. Phillips (Ed.), *High Energy Physics and Cosmology, 1998 Summer School*, 260–+

Liddle, A. R., and D. H. Lyth: 2000, *Cosmological Inflation and Large-Scale Structure*, Cambridge University Press, Cambridge, first edition

Lukash, V. N.: 2006, "LETTERS TO THE EDITORS: On the relation between tensor and scalar perturbation modes in Friedmann cosmology", *Physics Uspekhi* **49**, 103–106, arXiv:astro-ph/0610312

Lyth, D. H., and A. R. Liddle: 2009, *The Primordial Density Perturbation*, Cambridge University Press, Cambridge, first edition

Martin, J., and D. J. Schwarz: 1998a, "Influence of cosmological transitions on the evolution of density perturbations", *Phys. Rev. D* **57**, 3302–3316, arXiv:gr-qc/9704049

Martin, J., and D. J. Schwarz: 1998b, "Reply to 'Comment on "The influence of cosmological transitions on the evolution of density perturbations" '",, arXiv:gr-qc/9805069

Mather, J. C. et al.: 1994, "Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument", *Astrophys. J.* **420**, 439–444

McKay, D.: 2003, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, Cambridge, third edition

Miknaitis, G. et al.: 2007, "The ESSENCE Supernova Survey: Survey Optimization, Observations, and Supernova Photometry", *Astrophys. J.* **666**, 674–693, arXiv:astro-ph/0701043

Mukhanov, V. F., H. A. Feldman, and R. H. Brandenberger: 1992, "Theory of cosmological perturbations", *Phys. Rep.* **215**, 203–333

NASA: 2010, *Legacy Archive for Microwave Background Analysis (LAMBDA) website*, `http://lambda.gsfc.nasa.gov/`

Penzias, A. A., and R. W. Wilson: 1965, "A Measurement of Excess Antenna Temperature at 4080 Mc/s.", *Astrophys. J.* **142**, 419–421

Percival, W. J. et al.: 2010, "Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample", *Mon. Not. R. Astron. Soc.* **401**, 2148–2168, arXiv:0907.1660

Planck collaboration: 2005, "The scientific programme of Planck",, arXiv:0604069

Riess, A. G. et al.: 2009, "A Redetermination of the Hubble Constant with the Hubble Space Telescope from a Differential Distance Ladder", *Astrophys. J.* **699**, 539–563, arXiv:0905.0695

Samtleben, D., and QUIET collaboration: 2008, "QUIET - Measuring the CMB polarization with coherent detector arrays", *ArXiv e-prints*, arXiv:0806.4334

Smith, K. M. et al.: 2008, "CMBPol Mission Concept Study: Gravitational Lensing",, arXiv:0811.3916

Turner, M. S., and M. White: 1996, "Dependence of inflationary reconstruction upon cosmological parameters", *Phys. Rev. D* **53**, 6822–6828, arXiv:astro-ph/9512155

Zaldarriaga, M., and D. D. Harari: 1995, "Analytic approach to the polarization of the cosmic microwave background in flat and open universes", *Phys. Rev. D* **52**, 3276–3287, arXiv:astro-ph/9504085

Zaldarriaga, M., and U. Seljak: 1997, "All-sky analysis of polarization in the microwave background", *Phys. Rev. D* **55**, 1830–1840, arXiv:astro-ph/9609170

Zhao, W., D. Baskaran, and L. P. Grishchuk: 2009a, "On the road to discovery of relic gravitational waves: The TE and BB correlations in the cosmic microwave background radiation", *Phys. Rev. D* **79(2)**, 023002–+, arXiv:0810.0756

Zhao, W., D. Baskaran, and L. P. Grishchuk: 2009b, "Stable indications of relic gravitational waves in Wilkinson Microwave Anisotropy Probe data and forecasts for the Planck mission", *Phys. Rev. D* **80(8)**, 083005–+, arXiv:0907.1169